1 Brunn Minkowski Inequality

A set \( A \subseteq \mathbb{R}^n \) is called convex if \((1 - \lambda)x + \lambda y \in A\) for any \(x, y \in A\) and any \(\lambda \in [0, 1]\). The Minkowski sum of two sets \(A, B \subseteq \mathbb{R}^n\) is defined by

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

One sees that the sum of two convex sets remains convex, and that a set is convex if and only if \(\lambda K + (1 - \lambda)K = K\) for every \(\lambda \in (0, 1)\) (see exercise sheet).

The notion of convexity has been known since the ancient times. For example, it was noted by Archimedes that the area between a chord and a “convex” line is monotone with respect to changing the line, a property which clearly does not hold without the convexity assumption.

**Definition 1.1.** A convex body is a convex set \(K \subseteq \mathbb{R}^n\) which is compact and has non-empty interior. We say that \(K\) is centrally symmetric if \(K = -K\), that is, \(x \in K\) if and only in \(-x \in K\). We denote the class of convex bodies in \(\mathbb{R}^n\) by \(\mathcal{K}^n\).

Convex bodies are in correspondence with norms. If you did not do the following exercise in Chedva 2, do it now (see excercise sheet). We shall discuss convex sets in \(\mathbb{R}^n\), usually compact, usually with non-empty interior, not always centrally symmetric.

**Remark 1.2.** I am assuming you are all familiar with the notion of volume in \(\mathbb{R}^n\), which you can call dx or Lebesgue measure. It’s important to me here that it is additive, and that it is \(n\)-homogeneous, and that it can be approximated by unions of disjoint cubes. We in general like to discuss only measurable sets. One easily defines a convex set in \(\mathbb{R}^n\) which is not Borel measurable. It turns out, however, that convex sets are Lebesgue measurable. If you have no idea what Vol means: To define the volume of a set, take its indicator function, it should be integrable, and the value of this integral is the volume. Whether it is integrable of not depends on its boundary structure, namely that it is a null set, can be covered with boxes of \(\varepsilon\)-total volume for any \(\varepsilon > 0\). For convex sets this is true because of homothety. Later convex sets can be approximated by polytopes, and the volume of a polytope can be defined via pyramids, so we’ll have a more elementary notion of volume (which coincides with Lebesgue measure on convex sets).

A very basic inequality in AGA is the Brunn-Minkowski, which provides a fundamental relation between volume in \(\mathbb{R}^n\) and Minkowski addition.

**Theorem 1.3 (Brunn-Minkowski inequality).** Let \(K\) and \(T\) be two non-empty compact subsets of \(\mathbb{R}^n\). Then,

\[
(1) \quad \text{Vol}_n(K + T)^{1/n} \geq \text{Vol}_n(K)^{1/n} + \text{Vol}_n(T)^{1/n}.
\]

If we make the additional hypothesis that \(K\) and \(T\) are convex bodies, then we can have equality in (1) only if \(K\) and \(T\) are homothetical.

We will provide many different proofs for this inequality, each of which is beautiful and reflects different ideas. The inequality expresses in a sense the fact that volume, to the correct power, is a concave function with respect to Minkowski addition. For this reason, it is often written in the following form: if \(K\) and \(T\) are non-empty compact subsets of \(\mathbb{R}^n\) and \(\lambda \in (0, 1)\), then

\[
(2) \quad \text{Vol}_n(\lambda K + (1 - \lambda)T)^{1/n} \geq \lambda \text{Vol}_n(K)^{1/n} + (1 - \lambda)\text{Vol}_n(T)^{1/n}.
\]
Using (2) and the arithmetic-geometric means inequality we can write

\[ \text{Vol}_n(\lambda K + (1 - \lambda)T) \geq \text{Vol}_n(K)\lambda \text{Vol}_n(T)^{1-\lambda}. \]

This weaker form of the Brunn-Minkowski inequality has the advantage of being dimension-free. It is actually equivalent to (1) in the following sense: if we know that (3) holds for all \( K, T \) and \( \lambda \), we can recover (1) as follows.

Consider non-empty compact sets \( K \) and \( T \) (we may assume that both \( \text{Vol}_n(K) > 0 \) and \( \text{Vol}_n(T) > 0 \), otherwise there is nothing to prove), and define

\[
K_1 = \text{Vol}_n(K)^{-1/n}K, \quad T_1 = \text{Vol}_n(T)^{-1/n}T \quad \text{and} \quad \lambda = \frac{\text{Vol}_n(K)^{1/n}}{\text{Vol}_n(K)^{1/n} + \text{Vol}_n(T)^{1/n}}.
\]

Then, \( K_1 \) and \( T_1 \) have volume 1, and hence (3) gives

\[ \text{Vol}_n(\lambda K_1 + (1 - \lambda)T_1) \geq 1. \]

Since

\[ \lambda K_1 + (1 - \lambda)T_1 = \frac{K + T}{\text{Vol}_n(K)^{1/n} + \text{Vol}_n(T)^{1/n}}, \]

we immediately get (1).

**Remark 1.4** (Future teasers). *One can thinks of Brunn-Minkowski as (log-)concavity of the volume function with respect to Minkowski addition. Note, for example, that one can write

\[ \text{Vol}(A + tB) \geq \left(\text{Vol}(A)^{1/n} + t\text{Vol}(B)^{1/n}\right)^n \]

and there is equality for \( t = 0 \), which implies that there is an inequality of the derivatives with respect to \( t \). Note for later the equality when \( A = B \). Writing this derivative

\[ \lim_{t \to 0^+} \frac{\text{Vol}(A + tB) - \text{Vol}(A)}{t} \geq n\text{Vol}(B)^{1/n}\text{Vol}(A)^{n-1}. \]

The element in the left hand side is (OK, here I am being informal, use your intuition for this) “surface area” with respect to \( B \), and is usual surface area when \( B = B_2^n \) is Euclidean. Assume that \( A \) has the same volume as the Euclidean ball \( B_2^n \), we get

\[ \text{S.A.}(A) \geq n\text{Vol}(B_2^n) = S(B_2^n). \]

this is the isoperimetric inequality. (You did not have to remember surface area of \( B_2^n \), just note that in the case \( A = B_2^n \) there is equality all the way). In fact, the left hand side we shall see further in the course is, for convex sets, a polynomial in \( t \) by Steiner formula, which we shall later learn, namely that in the convex world

\[ \text{Vol}(K + tL) = \sum_{j=0}^{n} V_j(K, L)t^j \]

where \( V_0 = V(K) \) we get an inequalities between polynomials with equal leading coefficients, which implies an inequality of the second terms, namely

\[ V_1(K, L) \geq n\text{Vol}(K)^{n-1}\text{Vol}(L)^{1/n}. \]
this is called Minkowski’s inequality.

1.1 Proof using boxes

The first proof we discuss is due to Hadwiger and Ohmann. The argument is based on elementary sets. An elementary set is a finite union of non-overlapping boxes whose edges are parallel to the coordinate axes.

**Theorem 1.5.** Let \( A \) and \( B \) be elementary sets in \( \mathbb{R}^n \). Then,

\[
\Vol_n(A + B)^{1/n} \geq \Vol_n(A)^{1/n} + \Vol_n(B)^{1/n}.
\]

**Proof.** We first examine the case where both \( A \) and \( B \) are boxes. Assume that \( a_1, \ldots, a_n > 0 \) are the lengths of the edges of \( A \), and \( b_1, \ldots, b_n > 0 \) are the lengths of the edges of \( B \). Then, \( A + B \) is also a box, whose edges have lengths \( a_1 + b_1, \ldots, a_n + b_n \). Thus, we have to show that

\[
((a_1 + b_1) \cdots (a_n + b_n))^{1/n} \geq (a_1 \cdots a_n)^{1/n} + (b_1 \cdots b_n)^{1/n}.
\]

This is equivalent to the inequality

\[
\left( \frac{a_1}{a_1 + b_1} \cdots \frac{a_n}{a_n + b_n} \right)^{1/n} + \left( \frac{b_1}{a_1 + b_1} \cdots \frac{b_n}{a_n + b_n} \right)^{1/n} \leq 1.
\]

But, the arithmetic-geometric means inequality shows that the left hand side is less than or equal to

\[
\frac{1}{n} \left( \frac{a_1}{a_1 + b_1} + \cdots + \frac{a_n}{a_n + b_n} \right) + \frac{1}{n} \left( \frac{b_1}{a_1 + b_1} + \cdots + \frac{b_n}{a_n + b_n} \right) = 1.
\]

For every pair of elementary sets \( A \) and \( B \), we define the complexity of \( (A, B) \) as the total number of boxes in \( A \) and \( B \). We will prove the theorem by induction on the complexity \( m \) of \( (A, B) \). The case \( m = 2 \) was our first step.

Assume then that \( m \geq 3 \) and that the statement holds true for all pairs of elementary sets with complexity less than or equal to \( m - 1 \). Since \( m \geq 3 \), we may assume that \( A \) consists of at least two boxes. Let \( I_1 \) and \( I_2 \) be two of them. Since they are non-overlapping we can separate them by a coordinate hyperplane which, without loss of generality, may be described by the equation \( x_n = \rho \) for some \( \rho \in \mathbb{R} \). We define

\[
A^+ = A \cap \{ x \in \mathbb{R}^n : x_n \geq \rho \} \quad \text{and} \quad A^- = A \cap \{ x \in \mathbb{R}^n : x_n \leq \rho \}.
\]

Then, \( A^+ \) and \( A^- \) are non-overlapping elementary sets, and each one of them consists of a smaller number of boxes than \( A \). We now pass to \( B \), and find a hyperplane \( x_n = s \) such that if \( B^+ = B \cap \{ x \in \mathbb{R}^n : x_n \geq s \} \) and \( B^- = B \cap \{ x \in \mathbb{R}^n : x_n \leq s \} \), then

\[
(4) \quad \frac{\Vol_n(A^+)}{\Vol_n(A)} = \frac{\Vol_n(B^+)}{\Vol_n(B)} =: \lambda.
\]

Then, \( B^+ \) and \( B^- \) are elementary sets, we clearly have \( 0 < \lambda < 1 \), and the complexity of each pair \( (A^\pm, B^\pm) \) is less than \( m \). It is clear that

\[
A + B = (A^+ + B^+) \cup (A^+ + B^-) \cup (A^- + B^+) \cup (A^- + B^-).
\]

On the other hand, since \( A^+ + B^+ \subseteq \{ x : x_n \geq \rho + s \} \) and \( A^- + B^- \subseteq \{ x : x_n \leq \rho + s \} \), the
sets $A^+ + B^+$ and $A^- + B^-$ have disjoint interiors. Therefore,

$$\text{Vol}_n(A + B) \geq \text{Vol}_n(A^+ + B^+) + \text{Vol}_n(A^- + B^-).$$

Our inductive hypothesis can be applied to the right hand side, giving

$$\text{Vol}_n(A^+ + B^+) \geq \text{Vol}_n(A^+) + \text{Vol}_n(B^+)$$

and

$$\text{Vol}_n(A^- + B^-) \geq \text{Vol}_n(A^-) + \text{Vol}_n(B^-).$$

Taking the definition of $\lambda$ into account we get

$$\text{Vol}_n(A + B) \geq \lambda (\text{Vol}_n(A) + \text{Vol}_n(B))^n + (1 - \lambda) (\text{Vol}_n(A) + \text{Vol}_n(B))^n$$

which proves the theorem.

Since every compact set in $\mathbb{R}^n$ can be approximated by a decreasing sequence of elementary sets in the Hausdorff metric, Theorem 1.5 is easily extended to non-empty compact sets $K$ and $T$ in $\mathbb{R}^n$, by this proving Theorem 1.3 without the equality case.

### 1.2 Distance in $K^n$, Blaschke Selection theorem

Let us discuss “step back” and discuss a bit more the space of convex bodies, define the support function of a body, the distance between two bodies, and state and prove a useful theorem about compactness.

We shall endow the space of convex bodies in $\mathbb{R}^n$ with a natural topology. It can be introduced as the topology induced by the Hausdorff metric $\delta^H$:

**Definition 1.6.** For $K, T \in K^n$ we define

$$\delta^H(K, T) = \max \left\{ \max_{x \in K} \min_{y \in T} |x - y|, \max_{x \in T} \min_{y \in K} |x - y| \right\},$$

where $|x|$ denotes the Euclidean norm of $x \in \mathbb{R}^n$.

Put in another way,

$$\delta^H(K, T) = \inf\{\delta \geq 0 : K \subseteq T + \delta B_2^n, T \subseteq K + \delta B_2^n\},$$

where $B_2^n$ is the Euclidean unit ball.

Another way to look at convex bodies and at the Hausdorff distance is via the support function.

**Definition 1.7.** Given $K \subseteq \mathbb{R}^n$ non-empty and convex, we define the support function corresponding to $K$ by

$$h_K(u) = \sup_{x \in K} \langle x, u \rangle$$

where $u \in \mathbb{R}^n$. 

---

4
One may check that $h_K$ is 1-positively homogeneous, lower semicontinuous and convex. Oppositely one can show that every positively homogeneous lower semicontinuous convex $h : \mathbb{R}^n \to (-\infty, \infty]$ is the support function of some (unique) closed convex set $K$. The operation $K \mapsto h_K$ is order preserving: $K \subseteq T$ implies $h_K \leq h_T$ and $h_K \leq h_T$ implies $K \subseteq T$. The quantity $h_K(u) + h_K(-u)$ is called the width of $K$ in direction $u$.

A simple but important property of the support function regards the way it interacts with Minkowski addition. For $\lambda, \mu \geq 0$ one has for every $u \in \mathbb{R}^n$

$$h_{\lambda K + \mu T}(u) = \lambda h_K(u) + \mu h_T(u).$$

One has $h_{-K}(u) = h_K(-u)$, and so $K$ is centrally symmetric if and only if $h_K$ is even, while $0 \in K$ if and only if $h_K \geq 0$.

Our next observation is that

$$\delta^H(K, T) = \max\{|h_K(u) - h_T(u)| : u \in S^{n-1}\}.$$

Indeed, $K \subseteq T + \delta B^n_2$ if and only if $h_K \leq h_T + \delta$ on $S^{n-1}$. Thus, the embedding $K \mapsto h_K$ from $K^n$ to the space $C(S^{n-1})$ of continuous functions on the sphere is an isometry between $(K^n, \delta^H)$ and a subset of $C(S^{n-1})$ endowed with the supremum norm. Note that this mapping is positively linear (mapping Minkowski addition to sum of functions) and order-preserving (between inclusion and point-wise inequality). Note that there is no linearity for negative scalars. We mention a useful corollary from these simple facts.

**Corollary 1.8.** Given $K, L, M \in \mathcal{K}^n$ we have

$$\delta^H(K, L) = \delta^H(K + M, L + M).$$

In particular, we have the cancellation law

$$K + M = L + M \text{ implies } K = L.$$

Finally, we state and sketch the proof of a useful theorem of Blaschke

**Theorem 1.9.** Let $K_m \subset \mathbb{R}^n$ be a sequence of convex sets and assume for some $R > 0$ that $K_m \subset RB^n_2$ for all $m$. Then there exists a subsequence $K_{m_j}$ and a convex $K \subset \mathbb{R}^n$ with $K_{m_j} \to j \to \infty K$.

**Proof.** If you know the lemma of Ascoli and Arzela: Given a sequence of functions which are bounded uniformly ($\exists M > 0$ s.t. $|f_m(t)| \leq M$ for all $t \in S^{n-1}$ and $m$) and continuous uniformly (for any $\varepsilon > 0$ there exists $\delta > 0$ such that for any $s, t \in S^{n-1}, m \in \mathbb{N}$ with $|s - t| < \delta$ one have $|f_m(s) - f_m(t)| < \varepsilon$), there exists a convergent subsequence of $f_m$. You may then use it for $h_{K_m}$ which is uniformly bounded of the sphere by the assumption (with $R$ as a bound) and uniformly continuous since

$$|h_K(x) - h_K(y)| \leq |h_K(x - y)| \leq R|x - y|$$

and you are done, you have a converging subsequence. The fact that it corresponds to a body which is the limit is immediate.

If you do not know it, you may do it by hand as follows: Cover $RB^n_2$ with a cube of sidelength $2R$. Take families of subcubes dividing it, with sidelengths converging to 0. At each stage, an infinite sequence of the “current” sequence of bodies will touch exactly the same subset of small
rectangles (since there are a finite number of rectangles, hence a finite number of possibilities which of them to touch). Inductively, one may construct a sequence $K_{i,j}$ such that $K_{i,j}$ is a subsequence of $K_{i-j,j}$ (and of the original $K_j$) and the distance between any two elements in $K_{i,j}$ is at most $2^{-i}$. Now, take $K_{m,j} = K_{i,j}$. It must be a subsequence of $K_j$, and it is a Cauchy sequence. In fact, $K_{m,j} + 2^{-(j-1)}B_2^n$ is monotone decreasing since $K_{m,j+1} \subseteq K_{m,j} + 2^{-j}B_2^n$. So one may take the intersection of all of these, $K$ to get the actual limit, and we claim that this is also the limit of $K_{m,j}$. Clearly it is included in $K_{m,j} + \epsilon B_2^n$ for large enough $j$. On the other hand, taking $K + \epsilon B_2^n$ and $G$ as its interior, the sequence $K_{m,j} + 2^{-(j-1)} \setminus G$ this is again decreasing, so its limit is its intersection, but this intersection in now empty by definition of $K$, now use the finite intersection property to say that some finite intersection must be empty, that is, for large enough $j$ we have $K_j \subset G \subset K + \epsilon B_2^n$ as needed. ☐
2 Steiner symmetrization

If not done before: we shall denote $\kappa_n = \text{Vol}_n(B^n_2)$.

Remark 2.1. It is useful to know how $\kappa_n$ behaves as a function of $n$. It is not relevant to this chapter, but anyhow a good place to remark upon it. In Chedva 3 one shows
\[
(2\pi)^{n/2} = \int_{\mathbb{R}^n} e^{-x^2} \, dx = \int_{S^{n-1}} \int_0^\infty e^{-r^2} r^{n-1} n\kappa_n \, dr \, d\sigma = \int_0^\infty e^{-s} (2s)^{n/2 - 1} n\kappa_n \, dr = n\kappa_n 2^{n/2} \Gamma\left(\frac{n}{2}\right).
\]
so that $\kappa_n = \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)} = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)}$. In particular $\kappa_1 = 2$, $\kappa_2 = \pi$, $\kappa_3 = \frac{4}{3}\pi$, $\kappa_4 = \frac{\pi^2}{2}$ etc. Seems like a growing sequence? Not the case at all... Using the approximation of Stirling $\Gamma(k+1) \simeq \sqrt{2\pi k(k+1)}$ (where $\simeq$ means that the quotient converges to 1) we get $\kappa_n \simeq \frac{1}{\sqrt{\pi n}} \left(\frac{2\pi e}{n}\right)^{n/2}$. So, a ball of volume one has to have radius of the order $\sqrt{n}$.

2.1 Defining Steiner symmetrization, simple properties

An important process is that of successive Steiner symmetrizations that one can apply to a convex body so as to turn it into a ball. This method is very useful in proving geometric inequalities, as we shall see below.

Definition 2.2. Let $K \subseteq \mathbb{R}^n$ be a convex body and $u \in S^{n-1}$. The Steiner symmetrical $S_u(K)$ of $K$ in the direction of $u$ is defined so that for any $x \in u^\perp$
\[
\text{Vol}((x + \mathbb{R}u) \cap K) = \text{Vol}((x + \mathbb{R}u) \cap S_u(K)),
\]
and so that the right hand term is an interval centred at $x$.

The most useful fact about Steiner symmetrization, which was observed by Brunn, is that it preserves convexity.

Proposition 2.3. Let $K \subseteq \mathbb{R}^n$ be a convex body and $u \in S^{n-1}$, then $S_u(K)$ is convex, and has the same volume as $K$.

Proof. Convexity follows easily by considering trapezoids. Indeed, convexity is a two-dimensional notion, so it is enough to consider $x$ and $y$ in $u^\perp$ and check that for any $0 < \lambda < 1$ the interval centred at $z = (1 - \lambda)x + \lambda y$ with length $l_z = \text{Vol}(K \cap (z + \mathbb{R}u))$ has length greater than or equal to $(1 - \lambda)l_x + \lambda l_y$; the latter follows from the fact that, by convexity, $(1 - \lambda)(K \cap (x + \mathbb{R}u)) + \lambda (K \cap (y + \mathbb{R}u)) \subseteq K \cap (z + \mathbb{R}u)$. For the volume equality, simply use Fubini to compute the volume by integration on lines $x + u$ with $x \in u^\perp$. The lengths of these fibres remains the same after symmetrization.

Steiner symmetrization has many useful properties, many of which are left as an exercise. Clearly $S_u(x + \lambda K) = S_u(x) + \lambda S_u(K)$. For future reference we define some quantities associated with a convex body.

Definition 2.4. Let $K \subseteq \mathbb{R}^n$ be a convex set. We define the diameter of $K$ to be
\[
\text{diam}(K) = \sup\{|x - y| : x \in K, y \in K\} \in [0, \infty]
\]
The cricumradius of $K$ to be

$$R(K) = \inf \{ R : \exists x \in \mathbb{R}^n, x + RB_2^n \supseteq K \} \in [0, \infty]$$

The inradius of $K$ to be

$$r(K) = \sup \{ r : \exists x \in \mathbb{R}^n, x + rB_2^n \subseteq K \} \in [0, \infty]$$

For a centrally symmetric body, $\delta^H(K, \rho B_2^n) = R(K) - r(K)$ for $\rho \in [r(K), R(K)]$. In the exercise sheet you will show that $d, r$ and $R$ are monotone with respect to Steiner symmetrization. You will show that if $K_m \to K$ where all of them (including the limit) are convex bodies with non-empty interior, then $S_u K_m \to S_u K$. This is not assured without the assumption that int$(K) \neq \emptyset$.

Remarks 2.5. We will later show more exciting properties of $S_u$, in particular that $\partial(S_u(K)) \leq \partial(K)$, where $\partial(C)$ denotes the surface area of $C$, that $\Vol_n(K^\circ) \leq \Vol_n((S_u(K))^\circ)$ where $A^\circ$ denote the polar of $A$ (that is, the unit ball of $h_K$, we’ll get there soon) and that quermassintergrals in general are monotone $W_i(S_u(K)) \leq W_i(K)$ for all $i$ (these are the coefficients of the polynomial $\Vol(K + tB_2^n)$).

2.2 Convergence to a ball

Clearly $S_u(K)$ is “more symmetric than $K$” in that it is symmetric with respect to reflections around $u$ regardless of whether $K$ was. We can symmetrize with respect to the unit vector basis and get a body which is symmetric with respect to all coordinate reflections (this kind of body is called “unconditional”). When we sucessively reflect without orthogonality, however, we do not retain the symmetry with respect to the “older” reflections, which is why we should work a little bit in order to prove the following intuitive theorem.

Theorem 2.6. Let $K \subset \mathbb{R}^n$ be a convex body (with non-empty interior) such that $\Vol_n(K) = \Vol_n(B_2^n) =: \kappa_n$. Then there exists a sequence of vectors $u_j \in S^{n-1}$ such that, setting $K_j = S_{u_j}(K_{j-1})$, we have $K_j \to B_2^n$ (in the Hausdorff metric).

Proof. Consider the class $\mathcal{S}$ of all bodies obtained by successive symmetrizations of the original body $K$. They are all included in any ball around 0 that contains $K$. Denote

$$R_0 = \inf \{ R(T) : T \in \mathcal{S} \}.$$

(the infimal circumradius of bodies in this family). Take a sequence of symmetrals whose circumradius converges to $R_0$, and by Theorem 1.9 a convergent subsequence of it, $K_j \to L$. By continuity of circumradius, $R(L) = R_0$. First we show that the limit body $L$ is a ball. Assume not, then $L$ is included in $R_0 B_2^n$ (centered at the origin without loss of generality) and misses some cap of this ball (say a cap of radius $\delta > 0$). By compactness one may cover the boundary of $R_0 B_2^n$ with finitely many mirror images of this cap with respect to a sequence of hyperplanes. We claim sucessively symmetrizing $L$ with respect to each of the hyperplanes will produce a body which is included in $(R_0 - \delta')B_2^n$. Indeed, every cap for which we do not have a contact point of $L$ and $B_2^n$, will not gain a new contact point after symmetrization. Therefore, once all caps are detached, we have a smaller circumradius.

Pick a body $K_m$ in the sequence that is, say, $\delta'/(100n)$ close to $L$. Using the continuity of Steiner symmetrization (which is part of the exercise sheet), we see that symmetrizing the
body $K_m$ with respect to all these hyperplanes will produce a body included in a ball with radius $(R_0 - \delta')/2 < R_0$, contradicting the definition of $R_0$.

We are almost done, as we have shown that for any $\varepsilon$ we can find a sequence of symmetrizations which bring us close to a ball up to $\varepsilon$. We may now build our sequence inductively. Note that this proves in fact that we can find a sequence of symmetrizations with respect to which several bodies, together, converge to a ball. \hfill \qed

Remark 2.7. The question of how fast these symmetrizations approach $B^*_2$ (given the best choice of $u_j$, or a random choice, or a semi-random one) is of high importance in this field. Later we might explain why one needs approx $O(n)$ symmetrizations in order to approach a Euclidean ball within small distance.

Remark 2.8. There is a continuous way to look at Steiner symmetrization, letting the various intervals move in constant velocity (which depends on the interval). This is called a system of moving shadows, and we will possibly later go into this.

2.3 Again Brunn-Minkowski

We shall invoke the following simple property of Steiner symmetrizations which you will prove in the exercise sheet

Lemma 2.9. Let $K, T$ be convex sets in $\mathbb{R}^n$, and $u \in S^{n-1}$. Then $S_u(K) + S_u(T) \subseteq S_u(K + T)$.

Proof. Take $x \in S_u(K)$ and $y \in S_u(T)$, then $x = x' + t_x u$ and $y = y' + t_y u$ where $x', y' \in u^\perp$ and $|t_x| \leq \frac{1}{2} |(x' + Ru) \cap K|$, $|t_y| \leq \frac{1}{2} |(y' + Ru) \cap T|$. Their sum is $x' + y' + (t_x + t_y)u$ so all we need to explain is why

$$ |(x' + y' + Ru) \cap (K + T)| \geq |(x' + Ru) \cap K| + |(y' + Ru) \cap T|. $$

But this follows by the easily verified inclusion

$$(x' + y' + Ru) \cap (K + T) \supseteq (x' + Ru) \cap K + (y' + Ru) \cap T.$$

\hfill \qed

Proof of Brunn Minkowski again. Pick a sequence of Steiner symmetrizations $(S_{u_n})_{n \in \mathbb{N}}$ such that the symmetrals of $K$, of $T$, and of $K + T$ all converge to balls. By volume preservations, there balls radii, $r_K, r_T$ and $r_{K+T}$ respectively, satisfy

$$ r^n_K \kappa_n = \text{Vol}(K), r^n_T \kappa_n = \text{Vol}(T), r^n_{K+T} \kappa_n = \text{Vol}(K + T). $$

Note that by Lemma 2.9

$$ S_{u_2}(S_{u_1}(K + T)) \supseteq S_{u_2}(S_{u_1}(K) + S_{u_1}(T)) \supseteq S_{u_2}(S_{u_1}(K)) + S_{u_2}(S_{u_1}(T)),$$

and by induction this inclusion remains valid for each iterate. Therefore in the limit we get

$$ r_{K+T} B^n_2 \supseteq (r_K + r_T) B^n_2,$$

which is equivalent to the inequality $r_{K+T} \geq r_K + r_T$. Rewriting using the explicit radii and cancelling the term $\kappa_n^{1/n}$ we get

$$ \text{Vol}(K + T)^{1/n} \geq \text{Vol}(K)^{1/n} + \text{Vol}(T)^{1/n}. $$
2.4 Brunn concavity principle

Historically, the first proof of the Brunn-Minkowski inequality was based on Brunn’s concavity principle:

**Theorem 2.10.** Let $K$ be a convex body in $\mathbb{R}^n$ and let $F$ be a $k$-dimensional subspace. Then, the function $f : F^\perp \rightarrow \mathbb{R}$ defined by

$$f(x) = \text{Vol}_k (K \cap (F + x))^{1/k}$$

is concave on its support.

The following construction appears in the proof, and will be later useful, which is why we introduce it as a definition with the name “Schwartz symmetrization” of $K$ with respect to a subspace $F$

**Definition 2.11.** Let $K \subseteq \mathbb{R}^n$ be a convex body and $F \in G_{n,k}$ a subspace of dimension $k$. The Schwartz symmetrel $S_F(K)$ of $K$ with respect to $F$ is defined so that for any $x \in F^\perp$

$$\text{Vol} ((x + F) \cap K) = \text{Vol} ((x + F) \cap S_F(K)),$$

and so that the right hand term is a Euclidean ball centred at $x$.

**Remark 2.12.** Clearly the Schwartz symmetrization with $k = 1$ is simply the Steiner symmetrization. As a small variation of the argument in Theorem 2.6 explained above, one can find a sequence of successive Steiner symmetrizations in directions $u \in F$ so that the limiting convex body is $S_F K$. Indeed, these symmetrizations do not change the volumes of slices $(x + F) \cap K$ and in the limit makes all these slices arbitrarily close to Euclidean balls.

**Proof of Theorem 2.10.** Using the above remark, we see that $S_F K$ is a limit of steiner symmetrals of $K$, and in particular is convex. Denote the radius of the ball $S_F(K) \cap (x + F)$ by $r(x)$. The convexity of $S_F(K)$ implies that the function $r : F^\perp \rightarrow \mathbb{R}$ is concave on its support, which by definition of $r(X)$ means that and this shows that $\text{Vol} ((x + F) \cap K)^{1/k}$ is concave.

**Proof, once more, of the Brunn-Minkowski inequality.** Brunn’s concavity principle implies the Brunn-Minkowski inequality for convex bodies as follows. If $K$ and $T$ are convex bodies in $\mathbb{R}^n$, we define

$$K_1 = K \times \{0\} \text{ and } T_1 = T \times \{1\}$$

in $\mathbb{R}^{n+1}$ and consider their convex hull $L$. If

$$L(t) = \{x \in \mathbb{R}^n : (x, t) \in L\} \quad (t \in [0, 1])$$

we easily check that $L(0) = K$, $L(1) = T$ and

$$L(1/2) = \frac{K + T}{2}.$$

Then, Brunn’s concavity principle for $F = \mathbb{R}^n$ shows that

$$\text{Vol}_n \left( \frac{K + T}{2} \right)^{1/n} \geq \frac{1}{2} \text{Vol}_n(K)^{1/n} + \frac{1}{2} \text{Vol}_n(T)^{1/n},$$
and (1) follows.

2.5 Isoperimetric (revisited), and Isodiametric

Recall that the surface area of a convex body is given by

\[ S.A.(K) := \lim_{\varepsilon \to 0^+} \frac{\text{Vol}_n(K + \varepsilon B_2^n) - \text{Vol}_n(K)}{\varepsilon}. \]

Later in the course we shall discuss this expression in more detail, and show that it is in fact the "mixed volume" of \( K \) and \( (n-1) \times B_2^n \), and that it coincides with the Lebesgue area of the boundary \( \partial K \). We can easily check that it too is monotone with respect to Steiner symmetrizations.

**Proposition 2.13.** Let \( K \) be convex body and \( u \in S^{n-1} \). Then \( S.A.(S_u(K)) \leq S.A.(K) \) (surface area)

**Proof.** We use the definition of surface area as a limit of volume differences, and the property which was in the exercises as follows:

\[
S.A.(S_uK) = \lim_{\varepsilon \to 0^+} \frac{\text{Vol}_n(S_uK + \varepsilon B_2^n) - \text{Vol}_n(S_uK)}{\varepsilon} \leq \lim_{\varepsilon \to 0^+} \frac{\text{Vol}_n(S_u(K + \varepsilon B_2^n)) - \text{Vol}_n(S_uK)}{\varepsilon} = S.A.(K).
\]

Applying a sequence of Steiner symmetrizations \( K_j = S_{u_j}K_{j-1} \) such that \( K_j \to r_K B_2^n \) we get another proof for the isoperimetric inequality (of all bodies of fixed volume, the euclidean ball has smallest surface area). In fact we proved a stronger fact: Of all bodies of fixed volume, the Euclidean ball has smallest \( \varepsilon \)-extension, for any \( \varepsilon > 0 \).

Similarly, as you have shown in the exercises that diameter is monotone with respect to Steiner symmetrizations, you get

**Theorem 2.14 (Isodiametric).** For a convex body \( K \subset \mathbb{R}^n \) one has that

\[ \text{Vol}(K) \leq \left( \frac{\text{diam}(K)}{2} \right)^n \kappa_n. \]

**Remark 2.15.** One can prove this inequality directly from Brunn-Minkowski, try!
3 Functions instead of bodies

3.1 $\alpha$-concave and log-concave functions

We shall need the following definition for an $\alpha$-concave function: (Note that I have altered the definition a little from the class last Wed)

**Definition 3.1.** Let $K$ be a convex set in $\mathbb{R}^n$ and let $f : K \to \mathbb{R}^+$. We say that $f$ is $\alpha$-concave for some $\alpha > 0$ if $f^\alpha$ is concave on $K$. Equivalently, if for $0 < \lambda < 1$ and $x_0, x_1 \in K$

$$f^\alpha((1 - \lambda)x_0 + \lambda x_1) \geq (1 - \lambda)f^\alpha(x_0) + \lambda f^\alpha(x_1)$$

This definition has a natural extension to $\alpha = 0$ as follows:

**Definition 3.2.** A function $f : K \to \mathbb{R}^+$ ($K$ is a convex body) is called log-concave (or 0-concave) if

$$f((1 - \lambda)x_0 + \lambda x_1) \geq f((1 - \lambda)x_0)f^\lambda(x_1),$$

that is, if $\log f$ is concave.

**Remark 3.3.** As you will check in the homework assignment, this definition’s natural extension to $\alpha \in \mathbb{R}$ is as follows: A function is called $\alpha$-concave for $\alpha < 0$ (the name is not so good, actually) if

$$f((1 - \lambda)x_0 + \lambda x_1) \geq ((1 - \lambda)f^\alpha(x_0) + \lambda f^\alpha(x_1))^{1/\alpha},$$

that is, if $f^\alpha$ is convex (since taking power $\alpha$ of both sides changes the inequality sign).

Log concave functions play an essential role in the theory of convex bodies, and are considered the natural extension of convex bodies in the functional world. If $f$ is log-concave we sometimes let $f = \exp(-\varphi)$ where $\varphi$ is a convex function, allowed to attain the value $+\infty$ (where $f = 0$). It is many times useful to assume that $f$ is upper semi continuous (or that $\varphi$ is lower semi continuous) which just means that as $\varphi$ approaches the boundaries of its domain (where it is $\neq +\infty$) it cannot “jump”. Convex functions are always continuous on the interior of their domain. We shall discuss more on these functions shortly.

We mention that convex functions are important in optimization, mainly since a local minimum is global. Indeed, if $x_0$ is a local minimum, say in a ball around it or radium $\rho$, then take any $y$ farther from $x_0$ and create

$$x = \frac{\rho}{|y - x_0|}y + (1 - \frac{\rho}{|y - x_0|})x_0$$

which satisfies $|x - x_0| = \rho$ thus $f(x) \geq f(x_0)$ but $f(x) \leq \frac{\rho}{|y - x_0|}f(y) + (1 - \frac{\rho}{|y - x_0|})f(x_0)$ and thus $f(x_0) \leq f(y)$.

3.2 Milman-Gromov proof of Brunn’s principle

First we describe yet another, alternative, proof of Theorem 2.10, which is due to Gromov and Milman. Recall that in our language, the Brunn principle is that the marginal of the indicator function of a convex body onto a $(n - k)$-dimensional subspace is $1/k$-concave.

The following lemma is a simple application of Hölder’s inequality (applied for a two-point probability space). Hölder inequality states that

Given a measure $\mu$ and two functions $a(t) \in L_p(\mu), b(t) \in L_q(\mu)$ where $\frac{1}{q} + \frac{1}{p} = 1$

then one has $\int a(t)b(t)d\mu \leq (\int a^p(t)d\mu)^{1/p}(\int b^q(t)d\mu)^{1/q}$.
Lemma 3.4. Let \( f, g : K \to \mathbb{R}^+ \). If \( f \) is \( 1/\alpha \)-concave and \( g \) is \( 1/\beta \)-concave, then \( fg \) is \( 1/(\alpha + \beta) \)-concave.

Proof. Let \( x_0, x_1 \in K \) and \( \lambda \in (0, 1) \). Apply Hölder’s inequality for the measure \((1-\lambda)\delta_{x_0} + \lambda \delta_{x_1}\) with \( p = (\alpha + \beta)/\alpha \) and \( q = (\alpha + \beta)/\beta \), we get that

\[
(1-\lambda)(f(x_0)g(x_0))^{\frac{1}{\alpha+\beta}} + \lambda(f(x_1)g(x_1))^{\frac{1}{\alpha+\beta}} \\
\leq \left( (1-\lambda)f(x_0)^{\frac{1}{\alpha}} + \lambda f(x_1)^{\frac{1}{\beta}} \right)^{\frac{\alpha}{\alpha+\beta}} \left( (1-\lambda)g(x_0)^{\frac{1}{\alpha}} + \lambda g(x_1)^{\frac{1}{\beta}} \right)^{\frac{\beta}{\alpha+\beta}}
\]

and by the \( \alpha \)-concavity of \( f \) and the \( \beta \)-concavity of \( g \) we have that the latter

\[
\leq (f((1-\lambda)x_0 + \lambda x_1))^{\frac{1}{\alpha+\beta}} (g((1-\lambda)x_0 + \lambda x_1))^{\frac{1}{\alpha+\beta}}.
\]

We will use the above lemma to show how a one dimensional marginal of a function affects its level of concavity. Given a continuous function \( f : K \to \mathbb{R}^+ \) where \( K \subset \mathbb{R}^n \) and a vector \( \theta \in S^{n-1} \), we define the marginal \( \pi_{\theta \perp} f : P_{\theta \perp} K \to \mathbb{R}^+ \) in the usual way. We may assume that \( f \) is 0 outside of \( K \) and simply write for \( y \in P_{\theta \perp} K \) that

\[
(\pi_{\theta \perp} f)(y) := \int_R f(y + t\theta) dt
\]

(or write \( \int_{y+R\theta} f \)). It is sometimes useful to consider, for \( y \in P_{\theta \perp} (K) \), the interval on which \( f \) is actually defined, namely \( I_y = \{ t \in \mathbb{R} : y + t\theta \in K \} \). From the convexity of \( K \) it follows that \( I_y \) is an interval.

The main claim in this section is

Lemma 3.5. Let \( K \) be a convex body in \( \mathbb{R}^n \) and \( \theta \in S^{n-1} \). If \( f : K \to \mathbb{R}^+ \) is \( 1/\alpha \)-concave, then \( \pi_{\theta \perp} f \) is \( 1/(1+\alpha) \)-concave.

Proof. Concavity of a function is a “two-dimensional” notion, so we may clearly assume that \( K \subset \mathbb{R}^2 \), in which case \( P_{\theta \perp}(K) \) is an interval. Let \( y_0, y_1 \in P_{\theta \perp}(K) \) and write \( I_{y_i} = [a_i, b_i], i = 0, 1 \). For every \( \lambda \in [0, 1] \) with we set \( y_\lambda = (1-\lambda)y_0 + \lambda y_1 \). Then,

\[
I_{y_\lambda} \supseteq [(1-\lambda)a_0 + \lambda a_1, (1-\lambda)b_0 + \lambda b_1].
\]

We define \( c_i \in I_{y_i} \) by the equations

\[
\int_{a_i}^{b_i} f(y_i + t\theta) dt = 2 \int_{a_i}^{c_i} f(y_i + t\theta) dt = 2 \int_{c_i}^{b_i} f(y_i + t\theta) dt.
\]

If \( K' \) is the convex hull of the intervals \([y_i + c_i\theta, y_i + b_i\theta], i = 0, 1 \) and \( K'' \) is the convex hull of the intervals \([y_i + a_i\theta, y_i + c_i\theta], i = 0, 1 \), we define \( f' = f|_{K'} \) and \( f'' = f|_{K''} \). By the definition of \( c_i \), if \( \pi_{\theta \perp} f' \) and \( \pi_{\theta \perp} f'' \) are \( 1/(1+\alpha) \)-concave, then \( \pi_{\theta \perp} f \) satisfies the \( (1+\alpha) \)-concavity condition at \( y_0, y_1 \) and \( y_\lambda \) for every \( \lambda \). Thus, we only need to prove that \( \pi_{\theta \perp} f' \) and \( \pi_{\theta \perp} f'' \) are \( 1/(1+\alpha) \)-concave.
We can repeat this reduction any number of times. For every \( n \geq 2 \) we define partitions \( a_i = t_{0,i} < t_{1,i} < \cdots < t_{n-1,i} < t_{n,i} = b_i \) of \([a_i, b_i] \) such that

\[
\int_{a_i}^{b_i} f(y_i + t \theta) dt = n \int_{t_{k-1,i}}^{t_{k,i}} f(y_i + t \theta) dt
\]

for all \( k \in \{1, \ldots, n\} \). The same observation as above shows that it suffices to check that \( \pi_{\theta^k} f^k \) is \( 1/(1 + \alpha) \)-concave for every \( k \), where \( f^k = f|_{K(k)} \) and \( K(k) \) is the convex hull of the intervals \([y_i + t_{k-1,i} \theta, y_i + t_{k,i} \theta]\).

Passing to the limit (draw a picture!) we see that we have to check the following infinitesimal claim:

Let \( t_i \in I_{y_i} \) and \( d_i > 0 \), \( i = 0, 1 \). Given \( \lambda \in (0, 1) \), set \( y_{\lambda} = (1 - \lambda)y_0 + \lambda y_1 \), \( t(\lambda) = (1 - \lambda)t_0 + \lambda t_1 \) and \( d(\lambda) = (1 - \lambda)d_0 + \lambda d_1 \). Then, the function \( \lambda \mapsto f(y_{\lambda} + t(\lambda) \theta) \cdot d(\lambda) \) is \( (1 + \alpha) \)-concave.

This claim follows from Lemma 3.4, since \( \lambda \mapsto f(y_{\lambda} + t(\lambda) \theta) \) is \( \alpha \)-concave and the linear function \( \lambda \mapsto d(\lambda) \) is \( 1 \)-concave.

\[\square\]

Another proof of Brunn’s concavity principle (Theorem 2.10). The indicator function of \( K \) is constant on \( K \), and hence it is \( 1/\alpha \)-concave for every \( \alpha > 0 \). We choose an orthonormal basis \( \{\theta_1, \ldots, \theta_k\} \) of \( F \) and perform successive projections in the directions of \( \theta_i \). Lemma 3.5 shows that the function \( x \mapsto \text{Vol}_k(K \cap (F + x)) \) is \( 1/(\alpha + k) \)-concave on \( P_{F \perp}(K) \), for every \( \alpha > 0 \). Taking the limit of the pointwise inequality which is satisfied for any \( \alpha > 0 \), we see that it is satisfied in the limit, which means that \( f(x) = \text{Vol}_k(K \cap (F + x))^{1/k} \) is concave.

\[\square\]

### 3.3 Prékopa-Leindler inequality

The inequality of Prékopa and Leindler is considered the “functional analogue” for the Brunn Minkowski inequality. It is the following statement.

**Theorem 3.6** (Prékopa-Leindler). Let \( f, g, h : \mathbb{R}^n \to \mathbb{R}^+ \) be measurable functions, and let \( \lambda \in (0, 1) \). We assume that \( f \) and \( g \) are integrable, and that for every \( x, y \in \mathbb{R}^n \)

\[
h((1 - \lambda)x + \lambda y) \geq f(x)^{1-\lambda} g(y)^{\lambda}.
\]

Then,

\[
\int_{\mathbb{R}^n} h \geq \left( \int_{\mathbb{R}^n} f \right)^{1-\lambda} \left( \int_{\mathbb{R}^n} g \right)^{\lambda}.
\]

Before the proof, let us make a few remarks. First, it is indeed a generalization for the Vrnuk Minkowski inequality. Indeed,

**Proof of the Brunn-Minkowski inequality using PL.** Let \( K \) and \( T \) be non-empty compact subsets of \( \mathbb{R}^n \), and \( \lambda \in (0, 1) \). We define \( f = 1_K \), \( g = 1_T \), and \( h = 1_{(1-\lambda)K + \lambda T} \), where \( 1_A \) denotes the indicator function of a set \( A \). It is easily checked that the assumptions of Theorem 3.6 are satisfied, therefore

\[
\text{Vol}_n((1 - \lambda)K + \lambda T) = \int_{\mathbb{R}^n} h \geq \left( \int_{\mathbb{R}^n} f \right)^{1-\lambda} \left( \int_{\mathbb{R}^n} g \right)^{\lambda} = \text{Vol}_n(K)^{1-\lambda} \text{Vol}_n(T)^{\lambda}.
\]

\[\square\]
Another important remark is that this inequality is *in a sense* opposite to Hölder’s inequality. Indeed, the function $h$ in the assumption of the theorem satisfies that

$$h(z) \geq f(z)^{(1-\lambda)} g(z)^{\lambda},$$

as from Hölder is follows that

$$\int_{\mathbb{R}^n} f(z)^{1-\lambda} g(z)^{\lambda} dz \leq \left( \int_{\mathbb{R}^n} f(x) dx \right)^{1-\lambda} \left( \int_{\mathbb{R}^n} g(x) dx \right)^{\lambda},$$

which is an inequality in the opposite direction (but for a function smaller than $h$).

A third remark before we go into the proof is that for $f = g = h$, the condition in the theorem is precisely that $f$ is log concave. Of course, for $f = g = h$ the conclusion is trivial, but you will use this in a slightly less trivial way in the homework.

**Proof of Theorem 3.6.** The proof goes by induction on the dimension $n$.

(a) We will give two proofs for the case $n = 1$ (the second one appears as a remark after the proof). We may assume that $f$ and $g$ are continuous and strictly positive.

We define $x, y : (0, 1) \to \mathbb{R}$ by the equations

$$\int_{-\infty}^{x(t)} f = t \int_{\mathbb{R}} f \quad \text{and} \quad \int_{-\infty}^{y(t)} g = t \int_{\mathbb{R}} g.$$

In view of our assumptions, $x$ and $y$ are differentiable, and for every $t \in (0, 1)$ we have

$$x'(t)f(x(t)) = \int_{\mathbb{R}} f \quad \text{and} \quad y'(t)g(y(t)) = \int_{\mathbb{R}} g.$$

We now define $z : (0, 1) \to \mathbb{R}$ by

$$z(t) = (1 - \lambda)x(t) + \lambda y(t).$$

Since $x$ and $y$ are strictly increasing, $z$ too is strictly increasing, and the arithmetic-geometric means inequality shows that

$$z'(t) = (1 - \lambda)x'(t) + \lambda y'(t) \geq (x'(t))^{1-\lambda}(y'(t))^\lambda.$$

Hence, we can estimate the integral of $h$ making the change of variables $s = z(t)$, as follows:

$$\int_{\mathbb{R}} h = \int_0^1 h(z(t))z'(t) dt$$

$$\geq \int_0^1 h((1 - \lambda)x(t) + \lambda y(t))(x'(t))^{1-\lambda}(y'(t))^\lambda dt$$

$$\geq \int_0^1 f^{1-\lambda}(x(t))g^\lambda(y(t)) \left( \frac{\int_{\mathbb{R}} f}{f(x(t))} \right)^{1-\lambda} \left( \frac{\int_{\mathbb{R}} g}{g(y(t))} \right)^\lambda dt$$

$$= \left( \int_{\mathbb{R}} f \right)^{1-\lambda} \left( \int_{\mathbb{R}} g \right)^\lambda.$$

(b) **Inductive step:** We assume that $n \geq 2$ and the assertion of the theorem has been proved in all dimensions $k \in \{1, \ldots, n - 1\}$. Let $f, g$ and $h$ be as in the theorem. For every $s \in \mathbb{R}$ we
define $h_s : \mathbb{R}^{n-1} \to \mathbb{R}^+$ setting $h_s(w) = h(w,s)$, and $f_s, g_s : \mathbb{R}^{n-1} \to \mathbb{R}^+$ in an analogous way. From (5) it follows that if $x, y \in \mathbb{R}^{n-1}$ and $s_0, s_1 \in \mathbb{R}$ then

$$h_{(1-\lambda)s_0 + \lambda s_1}((1-\lambda)x + \lambda y) \geq f_{s_0}(x)^{1-\lambda}g_{s_1}(y)^\lambda.$$  

Defining $H(s) = \int_{\mathbb{R}^{n-1}} h_s, F(s) = \int_{\mathbb{R}^{n-1}} f_s, G(s) = \int_{\mathbb{R}^{n-1}} g_s$ our inductive hypothesis gives

$$H((1-\lambda)s_0 + \lambda s_1) = \int_{\mathbb{R}^{n-1}} h_{(1-\lambda)s_0 + \lambda s_1} \geq \left(\int_{\mathbb{R}^{n-1}} f_{s_0}\right)^{1-\lambda} \left(\int_{\mathbb{R}^{n-1}} g_{s_1}\right)^\lambda$$

$$= F^{1-\lambda}(s_0)G^\lambda(s_1).$$

Applying the inductive hypothesis once again, this time with $n = 1$, to the functions $F, G$ and $H$, we get

$$\int_1 h = \int_1 H \geq \left(\int_1 F\right)^{1-\lambda} \left(\int_1 G\right)^\lambda = \left(\int_1 f\right)^{1-\lambda} \left(\int_1 g\right)^\lambda,$$

where we have used Fubini’s theorem as well. 

\[\Box\]

**Remark 3.7.** The proof above used induction, and something called “measure transportation” in dimension one. Let us indicate another proof for the one dimensional case. We need to use usual Brunn-Minkowski in dimension 1. For convex sets (intervals) this is trivial (and there is actual equality), for measurable sets it is easy: Given two non-empty, bounded, and measurable $X, Y \subset \mathbb{R}$ such that $\lambda X + (1-\lambda)Y$ is measurable, we claim that

$$\text{Vol}_1(\lambda X + (1-\lambda)Y) \geq \lambda\text{Vol}_1(X) + (1-\lambda)\text{Vol}_1(Y).$$

Of course, we could just use Brunn-Minkowski, but if we want to avoid this: work with compact sets, and the inequality will follow for non-compact by approximation. Write $X' = \lambda X$ and $Y' = (1-\lambda)Y$. Translate $X'$ and $Y'$ so that $X' \cap Y' = \{0\}$ and $X' \subset \mathbb{R}^-$ and $Y' \subset \mathbb{R}^+$. Then $X' \cup Y' \subset X' + Y'$ and $\text{Vol}_1(X' \cup Y') = \text{Vol}_1(X') + \text{Vol}_1(Y')$ and the one-dimensional Brunn-Minkowski is complete. Next, for the 1-dimensional Prekopa-Leindler, Let $0 < \lambda < 1$ and let $f, g, h : \mathbb{R} \to [0,\infty)$ be integrable, and satisfy that for all $x, y \in \mathbb{R}$

$$h((1-\lambda)x + \lambda y) \geq f(x)^{(1-\lambda)}g(y)^\lambda.$$

We may assume that the functions are bounded, since all integrals are approximated by truncations, say, and that $\sup f(x) = \sup g(x) = 1$ since the inequality is scaling invariant (multiply $f$ by $c_f$, $g$ by $c_g$ and $h$ by $c_f^{(1-\lambda)}c_g^\lambda$).

Notice that by the condition on the three functions we have an inclusion of the form

$$\lambda\{x : f(x) \geq t\} + (1-\lambda)\{y : g(y) \geq t\} \subset \{z : h(z) \geq t\}.$$

Therefore, using the one-dimensional Brunn-Minkowski inequality, we know that

$$\lambda\text{Vol}_1(\{x : f(x) \geq t\}) + (1-\lambda)\text{Vol}_1(\{y : g(y) \geq t\}) \leq \text{Vol}_1(\{z : h(z) \geq t\}).$$

We use the equality $\int_1 h = \int_0^\infty \text{Vol}_1(\{z : h(z) \geq t\})dt$ (for $h, g$ and $f$) to get

$$\int_1 h \geq (1-\lambda)\int_1 f + \lambda\int_1 g \geq \left(\int_1 f\right)^{(1-\lambda)} \left(\int_1 g\right)^\lambda.$$
3.4 Geometric interpretations

Given two functions $f, g : \mathbb{R}^n \to \mathbb{R}^+$, say log-concave, which is the smallest function $h$ which satisfies the conditions in Theorem 3.6? To this end, we move to the exponent, and talk about convex functions. Assume $\varphi, \psi : \mathbb{R}^n \to \mathbb{R} \cup \infty$. Consider the smallest function $\zeta$ which satisfies

$$\zeta((1 - \lambda)x + \lambda y) \leq (1 - \lambda)\varphi(x) + \lambda\psi(y).$$

We define it by

$$\varphi \square_\lambda \psi(z) = \inf_{(1 - \lambda)x + \lambda y = z} (1 - \lambda)\varphi(x) + \lambda\psi(y).$$

This is called the $\lambda$-inf-convolution of $\varphi$ and $\psi$. It is obtained by taking the epi-graphs of $\varphi$ and of $\psi$, which are convex (unbounded) sets in $\mathbb{R}^{n+1}$, and taking their Minkowski average (possibly with a closure). In particular, this is always a convex function. In this sense, one may think of the $\lambda$-inf-convolution as a kind of “Minkowski average” of the functions, and if volume is given by integral (of $\exp -\varphi$) then Prékopa Leindler is really just a version of Brunn Minkowski.

In the world of log-concave functions, this convolution becomes a “sup-convolution”,

$$f \star_\lambda g(z) = \sup_{z = (1 - \lambda)x + \lambda y} f^{1 - \lambda}(x)g^\lambda(y),$$

and when $f$ and $g$ are log-concave then so is $f \star_\lambda g$.

This is probably a good time to introduce the Legendre transform.

**Defintion 3.8.** Let $\varphi : \mathbb{R}^n \to (-\infty, +\infty]$ be convex and not $\equiv +\infty$ (proper).

$$\mathcal{L}\varphi(y) = \varphi^*(y) = \sup_{\{x \in \mathbb{R}^n\}} (\langle x, y \rangle - \varphi(x)).$$

One easily checks several key properties of this transform, used in many mathematical fields. First, even without any assumption on $\varphi$, the resulting function is the supremum of linear functions, and so it is convex and lower-semi-continuous (this just means that the epi-graph is closed). We mention that the resulting function is proper (not all $+\infty$) since there is some linear function below $\varphi$, say $\langle \cdot, y \rangle + c$ and thus $\varphi^*(y) \leq \sup \langle x, y \rangle - (\langle x, y \rangle + c) = -c$.

More interestingly, starting with a convex l.s.c. function $\varphi$ we have that $\varphi^{**} = \varphi$. (For a general function we get something called the “convex envelope” of $\varphi$, similar to the closed convex hull of a general set.)

**Lemma 3.9.** Let $\varphi : \mathbb{R}^n \to (-\infty, +\infty]$ (proper, that is, not everywhere $+\infty$) be a convex function which is lower semi continuous. Then

$$\varphi^{**} = \varphi.$$

**Proof.** We shall prove it by two inequalities. One side is quite easy $\varphi^{**} \leq \varphi$. Indeed, by the definition of $\mathcal{L}$ we have that $\varphi^*(z) \geq \langle x, z \rangle - \varphi(x)$ for every $x$ and $z$, and thus

$$\varphi^{**}(x) = \sup_z (\langle x, z \rangle - \varphi^*(z)) \leq \sup_z \varphi(x) = \varphi(x).$$

For the other direction, we use that $\varphi$ is the supremum of affine functions below it (this is “separation” which one always needs for claims of this sort.)

$$\varphi = \sup\{\langle \cdot, y \rangle - c : \langle \cdot, y \rangle - c \leq \varphi\}.$$
The latter condition, after moving terms from side to side, can be written as \( c \geq \varphi^*(y) \). So,

\[
\varphi = \sup\{ \langle \cdot, y \rangle - c : c \geq \varphi^*(y) \} \leq \sup_y (\langle \cdot, y \rangle - \varphi^*(y)) = \varphi^{**}.
\]

Finally, the reason I introduced here the Legendre transform was so as to tell you that this kind of “summation for convex functions” given by the inf-convolution is simply usual addition after application of Legendre transform, namely

\[
\varphi \Box \psi := \inf_{x+y=z} (\varphi(x) + \psi(y)) = \mathcal{L}(\varphi + \mathcal{L}\psi)
\]

and

\[
\varphi \Box \lambda \psi = \mathcal{L}((1 - \lambda)\varphi + \lambda \mathcal{L}\psi).
\]

Indeed,

\[
\mathcal{L}(\varphi \Box \psi)(y) = \sup_x \langle x, y \rangle - (\varphi \Box \psi)(x)
\]

\[
= \sup_x \langle x, y \rangle - \inf_{z+w=x} (\varphi(z) + \psi(w))
\]

\[
= \sup_{z,w} \langle z, y \rangle + \langle w, y \rangle - \varphi(z) - \psi(w) = \mathcal{L}\varphi(y) + \mathcal{L}\psi(y).
\]

It turns out that not only sup-convolution preserves log-concavity. Log-concave functions are well behaved with respect to usual convolution, which is again a consequence of the Prékopa Leindler inequality, as we next show.

**Lemma 3.10.** Let \( f, g : \mathbb{R}^n \to [0, +\infty) \) be log-concave functions. Then

\[
(f \ast g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y)dy
\]

is a log-concave function, and so is \( P_E f : E \to [0, \infty) \) defined for a subspace \( E \) by

\[
P_E f(x) := \int_{E^\perp} f(x+y)dy.
\]

**Proof.** Fix \( x \) and \( y \), let \( z = (1 - \lambda)x + \lambda y \) and define the following three functions

\[
F(a) = f(a)g(x - a), G(b) = f(b)g(y - b), H(c) = f(c)g(z - c).
\]

It is easy to check that log-concavity of \( f, g \) implies that

\[
H((1 - \lambda)a + \lambda b) \geq F(a)^{(1-\lambda)} G(b)\lambda.
\]

Thus, by Prekopa-Leindler

\[
\int H \geq (\int F)^{1-\lambda} (\int G)^\lambda.
\]

This can be translated into

\[
(f \ast g)(z) \geq ((f \ast g)(x))^{(1-\lambda)} ((f \ast g)(y))^\lambda.
\]
The second assertion follows from the first by approximation of \(1_E\), but can also be done directly, since for every fixed \(x\) and \(y\) we have, for \(z = (1 - \lambda)z' + \lambda z'' \in E^\perp\) that

\[
f(((1 - \lambda)x + \lambda y) + z) \geq f(x + z'(1-\lambda)f(y + z'')^\lambda
\]

and by Prekopa Leindler on \(E^\perp\) we get that

\[
P_E f(((1 - \lambda)x + \lambda y) = \int_E f((1 - \lambda)x + y + z)dz \geq (P_E f(x))^{(1-\lambda)}(P_E f(y))^{\lambda}.
\]

In particular, marginals of indicator functions are log-concave, but we know this already, they are actually \(1/k\) concave for \(k\) the dimension of \(E^\perp\). Another result which is worth mentioning (might be as an exercise sometimes) is that marginals of convex bodies (of arbitrary dimension) are dense in the space of log-concave measures (with respect to locally uniform convergence).

### 3.5 More Prékopa Leindler type inequalities

You will show in exercise 3 that a measure with log-concave density is log-concave. But if the density is \(\alpha\)-concave, the measure has a different power of concavity. In order to state the next result, we introduce some notation.

**Definition 3.11.** Let \(p \neq 0\) and \(\lambda \in (0, 1)\). For all \(x, y > 0\) we set

\[
M_p^\lambda(x, y) = ((1 - \lambda)x^p + \lambda y^p)^{1/p}.
\]

If \(x, y \geq 0\) and \(xy = 0\) we set \(M_p^\lambda(x, y) = 0\). Observe that

\[
\lim_{p \to 0^+} M_p^\lambda(x, y) = x^{1-\lambda}y^\lambda.
\]

**Remark 3.12.** By Hölder’s inequality for a two-point probability measure, if \(x, y, z, w \geq 0\), \(a, b, \gamma > 0\) and \(\frac{1}{a} + \frac{1}{b} = \frac{1}{\gamma}\), then

\[
M_a^\lambda(x, y) \cdot M_b^\lambda(z, w) \geq M_\gamma^\lambda(xz, yw).
\]

Indeed, the measure is \((1 - \lambda)\delta_0 + \lambda \delta_1\), say, and we consider two functions, \(f\) which is equal to \(x^{\gamma}\) at 0 and \(y^{\gamma}\) at 1, and the other which is equal to \(z^{\gamma}\) at 0 and \(w^{\gamma}\) at 1. Use Hölder with \(p = a/\gamma\) and \(q = b/\gamma\) so that

\[
\int (fg) \leq (\int f^p d\mu)^{1/p}(\int g^q d\mu)^{1/q}
\]

which can be rewritten as

\[
((1 - \lambda)(xz)^{\gamma} + \lambda y^{\gamma}) \leq ((1 - \lambda)x^a + \lambda y^a)^{\gamma/a} \left((1 - \lambda)z^b + \lambda w^b\right)^{\gamma/b}
\]

as needed.
Proposition 3.13. Let $f, g, h : \mathbb{R}^n \to \mathbb{R}^+$ be measurable functions, and let $p > 0$ and $\lambda \in (0, 1)$. We assume that $f$ and $g$ are integrable, and for every $x, y \in \mathbb{R}^n$

\begin{equation}
(6) \quad h((1 - \lambda)x + \lambda y) \geq M_p^\lambda(f(x), g(y)).
\end{equation}

Then,

$$
\int_{\mathbb{R}^n} h \geq M_p^\lambda \left( \int_{\mathbb{R}^n} f, \int_{\mathbb{R}^n} g \right).
$$

Proof. We will consider only the case $n = 1$. As in the proof of the Prékopa-Leindler inequality, we define $x, y : (0, 1) \to \mathbb{R}$ by the equations

$$
\int_{-\infty}^{x(t)} f = t \int f \quad \text{and} \quad \int_{-\infty}^{y(t)} g = t \int g.
$$

Then,

$$
x'(t)f(x(t)) = \int f \quad \text{and} \quad y'(t)g(y(t)) = \int g.
$$

We define $z : (0, 1) \to \mathbb{R}$ by

$$
z(t) = (1 - \lambda)x(t) + \lambda y(t).
$$

Then, $z$ is strictly increasing, and

$$
z'(t) = \lambda x'(t) + (1 - \lambda)y'(t) = M_p^\lambda(x'(t), y'(t)).
$$

Hence, using (3.12) and (6) we can estimate the integral of $h$ making the change of variables $s = z(t)$, as follows:

$$
\int_{\mathbb{R}} h = \int_0^1 h(z(t))z'(t)dt \\
\geq \int_0^1 h((1 - \lambda)x(t) + \lambda y(t))M_p^\lambda(x'(t), y'(t))dt \\
\geq \int_0^1 M_p^\lambda(f(x(t)), g(y(t)))M_p^\lambda(x'(t), y'(t))dt \\
\geq \int_0^1 M_p^\lambda \left( f(x(t))x'(t), g(y(t))y'(t) \right) dt \\
= M_p^\lambda \left( \int_{\mathbb{R}} f, \int_{\mathbb{R}} g \right).
$$

The inductive step is exactly as in Theorem 3.6.

Remark 3.14. Using Proposition 3.13 we may give an alternative proof of Lemma 3.4. The claim was the following: Assume that $K$ is a two-dimensional convex body and $f : K \to \mathbb{R}^+$ is a $1/\alpha$-concave function. If

$$
(Pf)(y) := \int_\mathbb{R} 1_K(y, t)f(y, t)dt,
$$

then $Pf$ is $1/(1 + \alpha)$-concave. Indeed, let $F_y(t) = 1_K(y, t)f(y, t)$, $y \in PK$. Then, for all $y, z \in PK$ we have

$$
F_{(1-\lambda)y+\lambda z}((1-\lambda)t+\lambda s) = 1_K((1-\lambda)(y, t)+\lambda(z, s))f((1-\lambda)(y, t)+\lambda(z, s)) \geq M_{1/\alpha}^\lambda(F_y(t), F_z(s))
$$

20
by the $1/\alpha$-concavity of $f$ and the convexity of $K$. The claim follows from Proposition 3.13 (the dimension is 1).

We close this section with one more functional inequality that will be used in the next section. It is a version of Prékopa Leindler with the condition on $h, f, g$ different - here $h$ is evaluated at the Harmonic $\lambda$-mean of $x$ and $y$, and is compared with a mean of $f$ and $g$ which is harmonic but with parameters depending on $x, y$ and $\lambda$. Of course, there is not “harmonic mean” of vectors, and we shall only discuss function on $\mathbb{R}^+$.

**Theorem 3.15.** Let $0 < \lambda < 1$ and $w, g, h : \mathbb{R}^+ \to \mathbb{R}^+$ be integrable functions such that

$$
(7) \quad h((1 - \lambda)r^{-1} + \lambda s^{-1}) \geq f(r)^{(1 - \lambda)\lambda r + \lambda s} g(s)^{(1 - \lambda)\lambda r + \lambda s}
$$

for every pair $(r, s) \in \mathbb{R}^+ \times \mathbb{R}^+$. Then,

$$
(8) \quad \int_0^\infty h \geq (1 - \lambda) \left( \int_0^\infty f \right)^{-1} + \lambda \left( \int_0^\infty g \right)^{-1}.
$$

**Proof.** We may assume that $f$ and $g$ are continuous and strictly positive. We define $r, s : [0, 1] \to \mathbb{R}^+$ by the equations

$$
\int_0^{r(t)} f = t \int_0^\infty f \quad \text{and} \quad \int_0^{s(t)} g = t \int_0^\infty g.
$$

Then, $r$ and $s$ are differentiable, and for every $t \in (0, 1)$ we have

$$
r'(t)f(r(t)) = \int_0^\infty f \quad \text{and} \quad s'(t)g(s(t)) = \int_0^\infty g.
$$

Next, we define $z : [0, 1] \to \mathbb{R}^+$ by

$$
z(t) = M_{\lambda, 1}(r(t), s(t)) = \frac{1}{(1 - \lambda)(r(t))^{-1} + \lambda(s(t))^{-1}} = \frac{r(t)s(t)}{(1 - \lambda)s(t) + \lambda r(t)}.
$$

Note that (writing $s = s(t), r = r(t)$)

$$
z'(t) = z^2(t) \left( (1 - \lambda)\frac{r'}{r^2} + \lambda\frac{s'}{s^2} \right)
$$

$$
= (1 - \lambda) \int \frac{f}{f(r)} \left( \frac{s}{(1 - \lambda)s + \lambda r} \right)^2 + \lambda \int \frac{g}{g(s)} \left( \frac{r}{(1 - \lambda)s + \lambda r} \right)^2
$$

$$
= (1 - \lambda) \int \frac{f}{f(r)} \left( \frac{s}{(1 - \lambda)s + \lambda r} \right)^{(1 - \lambda)\lambda r + \lambda s} \left( \frac{g}{g(s)} \right)^{(1 - \lambda)\lambda r + \lambda s} 

\geq \left( \int \frac{f}{f(r)} \left( \frac{s}{(1 - \lambda)s + \lambda r} \right)^{(1 - \lambda)\lambda r + \lambda s} \left( \frac{g}{g(s)} \right)^{(1 - \lambda)\lambda r + \lambda s} 
$$

by the arithmetic-geometric means inequality (with weights depending on $r$ and $s$, and $\lambda$).
Now, making a change of variables we write

\[
\int_{0}^{\infty} h = \int_{0}^{1} h(z(t))z'(t)dt \geq \int_{0}^{1} f(r)^{(1-\lambda)s} g(s)^{\lambda r} z'(t)dt \\
\geq \int_{0}^{1} \left( \int f \left( \frac{s}{(1-\lambda)s + \lambda r} \right)^{(1-\lambda)s} \right) g \left( \frac{r}{(1-\lambda)s + \lambda r} \right) dt \\
\geq \int_{0}^{1} \left( \frac{(1-\lambda)s}{(1-\lambda)s + \lambda r} \int f \left( \frac{s}{(1-\lambda)s + \lambda r} \right)^{-1} + \frac{\lambda r}{(1-\lambda)s + \lambda r} \int g \left( \frac{r}{(1-\lambda)s + \lambda r} \right)^{-1} \right) dt \\
= \left( 1 - \lambda \right) \left( \int_{0}^{\infty} f \right)^{-1} + \lambda \left( \int_{0}^{\infty} g \right)^{-1}. \]

This completes the proof. \qed
4 Blaschke-Santaló Inequality

4.1 The polar body

Definition 4.1. Let $K \subset \mathbb{R}^n$ be a convex set such that $0 \in K$. The polar of $K$ is defined by

$$K^\circ = \{ y \in \mathbb{R}^n : \langle x, y \rangle \leq 1 \forall x \in K \}.$$

Note that $K^\circ = \{ y : h_K(y) \leq 1 \}$, that is, $K^\circ$ is the unit ball corresponding to the norm given by $h_K$. Of course, this is not really necessarily a norm, since we did not assume $K$ was centrally symmetric, nor did we assume $K$ bounded, or that $0$ belongs to its interior. However, it can be called a “generalized norm”, which is allowed to attain $0$ and $+\infty$, and which is not required to satisfy $\| - x \| = \| x \|$, but merely positively-1-homogeneous, and convex; $\| x + y \| \leq \| x \| + \| y \|$. If we denote the norm with unit ball $K$ by $\| \cdot \|$, that is

$$\| x \|_K = \inf \{ r > 0 : x \in rK \},$$

and the dual norm by $h_K(\cdot) = \| \cdot \|_{K^\circ}$, we have by definition that for all $x$ and $y$

$$\langle x, y \rangle \leq \| x \|_K \cdot \| y \|_{K^\circ},$$

a useful property.

For polarity one can show a list of good properties (you will do this in the exercises): It interchanges order of inclusion, it is an involution on closed convex sets which include 0, and the set of compact convex sets which include 0 in the interior is invariant under this transformation, so that it is an involution there as well. It sends convex hull to intersection and vice versa. It is very much connected with the Legendre trasform $L$, as you will show too.

Lemma 4.2. The only set invariant under polarity is the euclidean unit ball.

Proof. Indeed, assume $K = K^\circ$ and denote by $\| \cdot \|$ the associated “generalized norm”. Apply inequality (9) for $x = y$ to get

$$|x|^2 = \langle x, x \rangle \leq \| x \|_K \cdot \| x \|_{K^\circ} = \| x \|^2$$

which means that $B_2^n \supseteq K$. But polarity changes order of inclusion so $B_2^n \subseteq K^\circ = K$ hence they are equal. \qed

4.2 Volume product

Of particular interest in the product of the volumes of $K$ and $K^\circ$. One defines the volume product of a centrally symmetric $K = -K$ by

$$s(K) = \text{Vol}_n(K)\text{Vol}_n(K^\circ).$$

Note that this number is invariant under invertible linear transformations. Indeed,

$$(AK)^\circ = \{ x : \langle Ay, x \rangle \leq 1 \forall y \in K \} = \{ x : \langle y, A^*x \rangle \leq 1 \forall y \in K \} = (A^*)^{-1}K^\circ.$$

Thus $s(AK) = \det(A) \det(A^*)^{-1}s(K) = s(K)$.

For non-symmetric convex bodies, one should be more careful. Since translation of a body changes the volume of the polar body, clearly the choice of where the origin is will affect the
Santalo product. For a body with 0 close to its boundary, the polar body will be huge, and its volume huge. Therefore for a general convex body one defines

\[ s(K) = \inf_x \{ \text{Vol}_n(K) \text{Vol}_n((K-x)^\circ) \}. \]

We shall later see (it is easy) that this infimum is attained, and that for centrally symmetric bodies it is attained at the origin, so that the two definitions coincide for \( K = -K \). The main theorem in this section is that \( s(K) \) is maximized when \( K \) is an ellipsoid. This is called the Blascke-Santaló inequality. The centrally symmetric case will be our first theorem. The non-symmetric case will be slightly more involved, and we will prove it by a reduction to the symmetric case.

The question of where the volume product is minimized remains to this day an open problem (even in dimension 3), and Mahler’s conjecture asserts that the minimizer is a simplex. Restricting to the class of centrally symmetric convex bodies, the conjecture asserts that the minimum is attained at the cube (however, in such a case it is far from unique). We shall discuss this in Section 4.6 below, together with asymptotic results. It is useful already to notice that

\[ s([−1,1]^n) = \text{Vol}([−1,1]^n) \times \text{Vol}(\text{conv}(±e_i)) = 2^n \times \frac{2^n}{n!} = \frac{4^n}{n!} \approx \frac{4^n}{\sqrt{2\pi n}(n/e)^n} = \frac{1}{\sqrt{2\pi n}} \left( \frac{4e}{n} \right)^n \]

That is, after taking power \( 1/n \), the upper bound, and the estimated lower bound, are very close to one another. It is a theorem of Bourgain and Milman that there exists a universal constant \( c > 0 \) (independent of dimension) such that

\[ c^n \kappa_n^2 \leq s(K) \leq \kappa_n^2 \]

for all \( K \). We might eventually get to proving this towards the end of the course.

### 4.3 Monotonicity with respect to Steiner

Our first proof, for the centrally symmetric case, uses Steiner symmetrizations to prove Santalo’s inequality. As usual, we shall show that symmetrization only increases the volume product.

**Theorem 4.3** (Increase dual volume). Let \( K = -K \) be a compact convex set with 0 in its interior, and let \( u \in S^{n-1} \). Then

\[ \text{Vol}(K^\circ) \leq \text{Vol}((S_u K)^\circ). \]

**Proof.** Without loss of generality we shall assume \( u = e_n \) and symmetrize with respect to the hyperplane \( x_n = 0 \). It will be useful to denote, for a set \( A \subset \mathbb{R}^n \), the \( (n-1) \) dimensional part of \( A \) at height \( r \) by \( A(r) \), that is,

\[ A(r) = \{ x \in \mathbb{R}^{n-1} : (x, r) \in A \}. \]
One checks directly that for $u = e_n$

$$S_u K = \{(x, \frac{1}{2}(s-t)) : (x, s), (x, t) \in K\},$$

and thus

$$(S_u K) = \{(y, r) : \langle x, y \rangle + r(s-t)/2 \leq 1 \ \forall (x, s), (x, t) \in K\}.$$ Consider for $r > 0$ the sets $K^\circ(r)$ and $K^\circ(-r)$. (If $K$ would have been symmetric about $x_n = 0$, then so would $K^\circ$ (check!) and these two sets would be equal.) We claim

$$\frac{1}{2} (K^\circ(r) + K^\circ(-r)) \subseteq (S_u K)^\circ(r).$$

Indeed,

$$\frac{1}{2} (K^\circ(r) + K^\circ(-r)) = \left\{ \frac{y+z}{2} : \langle x, y \rangle + sr \leq 1, \langle w, z \rangle - tr \leq 1 \ \forall (x, s), (w, t) \in K \right\}$$

$$\subseteq \left\{ \frac{y+z}{2} : \langle x, y \rangle + sr \leq 1, \langle x, z \rangle - tr \leq 1 \ \forall (x, s), (x, t) \in K \right\}$$

$$\subseteq \left\{ \frac{y+z}{2} : \frac{1}{2} \langle x, y+z \rangle + \frac{(s-t)}{2} r \leq 1 \ \forall (x, s), (x, t) \in K \right\}$$

$$= \left\{ u : \langle x, u \rangle + \frac{(s-t)}{2} r \leq 1 \ \forall (x, s), (x, t) \in K \right\}$$

= $$(S_u K)^\circ(r).$$

Since $K^\circ = -K^\circ$ is symmetric, we have we have

$$K^\circ(-r) = \{ x : (x, -r) \in K^\circ \} = \{ x : (-x, r) \in K^\circ \} = \{ -y : (y, r) \in K^\circ \} = -K^\circ(r),$$

and in particular both sets have the same volume. (Please notice that for a centrally symmetric $A$ there is no assurance that $A(r)$ is centrally symmetric.) By Brunn-Minkowski’s inequality we get that

$$\text{Vol}_{n-1}(\frac{K^\circ(r) + K^\circ(-r)}{2}) \geq \text{Vol}_{n-1}(K^\circ(r))^{1/2} \text{Vol}_{n-1}(K^\circ(-r))^{1/2} = \text{Vol}_{n-1}(K^\circ(r)).$$

Putting these together we see that

$$\text{Vol}_n(S_u(K))^\circ = \int_{-\infty}^{\infty} \text{Vol}_{n-1}((S_u(K))^\circ(r)) dr \geq \int_{-\infty}^{\infty} \text{Vol}_{n-1}(K^\circ(r)) dr = \text{Vol}_n(K).$$

Remark 4.4. Note that we have used central symmetry in an essential way. Check for yourselves that this proof does not carry over the non-centrally-symmetric case. One of the tricks which we shall use there would be to compare level sets with different $r$’s, not just $r$ with $-r$.

As a result we prove

Theorem 4.5 (Blaschke-Santaló). For any centrally symmetric convex body $K \in K_n^0$ we have

$$s(K) = \text{Vol}(K) \text{Vol}(K^\circ) \leq \kappa_n^2.$$
Proof. The inequality is clearly linear invariant (see the discussion after Definition 4.1) so we may assume without loss of generality that \( \text{Vol}(K) = \kappa_n \). Using Theorem 2.6 we take a sequence of subsequent symmetrizations \( K_j \) of \( K \) converging to \( B_n^2 \). Then by order reversion \( K_j \circ \rightarrow (B_n^2) \circ = B_n^2 \). Indeed,

\[
(1 - \varepsilon)B_n^2 \subset K_j \subset (1 + \varepsilon)B_n^2 \leftrightarrow (1 + \varepsilon)^{-1}B_n^2 \subset K_j \circ \subset (1 - \varepsilon)^{-1}B_n^2.
\]

Since Steiner symmetrization preserves volume, \( \text{Vol}_n(K_j) = \text{Vol}_n(K_j \circ) \), and by Theorem 4.3 \( \text{Vol}_n(K_j \circ) \geq \text{Vol}_n(K_{j-1} \circ) \), we see that the volume product is increasing,

\[
s(K) \leq s(K_1) \leq \cdots \leq s(K_{j-1}) \leq s(K_j) \rightarrow s(B_n^2) = \kappa_n^2.
\]

\[\square\]

Remark 4.6. It is used in the proof that if \( K \) is centrally symmetric then so is \( S_u K \), a fact which, if you did not check before, this is a good time to check.

4.4 Non-symmetric case

4.4.1 Preliminaries

To discuss the non-symmetric case let us begin with a few easy but illustrative facts. First, a formula for the volume of a convex body.

Lemma 4.7. For a compact convex body \( K \) with \( 0 \in \text{int}(K) \), we have the following formula

\[
\text{Vol}(K) = \kappa_n \int_{S^{n-1}} \|u\|_K^n d\sigma(u).
\]

Here \( \|x\|_K = \inf\{r : x \in rK\} \) is the gauge function generalized norm associated to \( K \).

Proof. Follows by polar integration. Using polar coordinates (the constant \( c_n \) can be computed directly, but for us it doesn’t matter at this point)

\[
\text{Vol}(K) = \int_{\mathbb{R}^n} 1_K(x) dx = \int_{S^{n-1}} \int_{0}^{\infty} 1_K(ru) r^{n-1} dr d\sigma(u) c_n.
\]

Clearly \( 1_K(ru) = 1 \) if and only if \( r \leq 1/\|u\|_K \), and is 0 otherwise, hence

\[
\text{Vol}(K) = \int_{S^{n-1}} \int_{0}^{1/\|u\|_K} r^{n-1} dr d\sigma(u) c_n = \int_{S^{n-1}} (1/\|u\|_K)^n d\sigma(u) c'_n.
\]

Finally to determine \( c'_n \) we simply plug in \( K = B_n^2 \) in which case \( \|u\|_K = 1 \) for \( u \in S^{n-1}, \) so that

\[c'_n = \kappa_n.\]

In the case of dual-volume, we get of course that

Lemma 4.8. For a compact convex body \( K \) with \( 0 \in \text{int}(K) \),

\[
\text{Vol}(K^\circ) = \kappa_n \int_{S^{n-1}} h_K(u)^{-n} d\sigma(u).
\]
Let us define, for \( z \in K \), a body which we shall call “dual with respect to \( z \)”, which is the dual of \( K \) in a world where the origin is at \( z \). In other words,

\[
K^z = (K - z)^0 + z.
\]

The extra \( +z \) is the for aesthetic reasons, namely so that (if \( z \in K \)) we have

\[
(K^z)^0 = ((K - z)^0 + z - z)^0 + z = K.
\]

**Lemma 4.9.** For a compact convex body \( K \) with non-empty interior, the function, defined on \( \text{int}(K) \), given by \( F(z) = \text{Vol}(K^z) \), is strictly convex in \( z \), and approaches \( +\infty \) as \( z \to \partial K \).

**Proof.** We have \( K \subseteq RB_2^n \) for some \( R \), so that \( K - z \subseteq 2rB_2^n \) for any \( z \in K \), therefore \((K - z)^0 \subseteq \frac{1}{2r}B_2^n\) for all \( z \in K \). As \( z \to \partial K \) we have that \( 0 \to \partial K - z \) and therefore the diameter of \((K - z)^0\) tends to \( \infty \), which, together with \((K - z)^0\) including a small ball, implies that \( \text{Vol}((K - z)^0) \) tends to \( \infty \) as well.

For the convexity part, we shall use Lemma 4.8. The support function of \((K - z)\) is simply \( h_K - \langle \cdot, z \rangle \), and thus

\[
F(z) = \int_{S^{n-1}} \left(h_K(u) - \langle u, z \rangle\right)^{-n} d\sigma(u).
\]

Each of the intergands is a strictly convex function of \( z \), and thus so is their integral. \( \square \)

We use these facts to show

**Proposition 4.10.** The Function \( F(z) = \text{Vol}(K^z) : \text{int}(K) \to \mathbb{R^+} \) has a unique minimum. (This point \( z_0 \) is called the Santaló point of \( K \).) For centrally symmetric \( K \) this minimum is attained at the origin. For general \( K \) it is attained on the unique point \( z_0 \) for which

\[
\int_{K^0 - z_0} x dx = 0.
\]

**Proof.** A strictly convex function has a unique minimum. If \( K \) is centrally symmetric, then \( K - z = (K + z)^0 - z = (K - z)^0 - z = -(K + z)^0 - z = -(K^z) \) and in particular the function \( F \) is even. An even convex function has minimum at 0. Finally, in general to find the minimum we differentiate under the integral sign

\[
\nabla_z F(z) = \int_{S^{n-1}} \nabla_z \left(h_K(u) - \langle u, z \rangle\right)^{-n} d\sigma(u) = n \int_{S^{n-1}} u \left(h_K(u) - \langle u, z \rangle\right)^{-(n+1)} d\sigma(u)
\]

At the point \( z_0 \) this gradient must be 0, that is,

\[
0 = \int_{S^{n-1}} \frac{u}{\|u\|^n (K - z)^0} d\sigma(u) = \int_{S^{n-1}} c_n \int_0^{1/\|u\|(K - z)^0} r^{n-1} d\sigma(u) = d_n \int_{\mathbb{R^n}} x 1_{(K - z)^0}(x) dx.
\]

\( \square \)

### 4.4.2 A key Lemma

The key Lemma for the non-symmetric case, proven by Meyer and Pajor, is that given a convex body \( K \), for every affine hyperplane \( H \) intersecting \( K \), and a point \( z \in K \cap H \), if we let \( \lambda \) denote the ratio into which \( H \) separates \( K^z \), that is, \( \text{Vol}(H^+ \cap K^z) = \lambda \text{Vol}(K^z) \)
and \( \text{Vol}(H^{-} \cap K^{z})(1 - \lambda)\text{Vol}(K^{z}) \), then Steiner symmetrizing \( K \) with respect to \( H = u^{\perp} + z \) (keeping \( z \) intact) gives a body \( S_{u}K \) with

\[
\text{Vol}((S_{u}K)^{z}) \geq 4\lambda(1 - \lambda)\text{Vol}(K^{z}).
\]

In particular, if \( H \) is medial to \( K^{z} \) then the volume of the \( z \)-dual is increased.

For simplicity, let is consider only the case \( \lambda = 1/2 \). The general case is treated very similarly. We thus prove:

**Lemma 4.11.** Assume \( K \subset \mathbb{R}^{n} \) is a convex body and \( H = u^{\perp} \) is a hyperplane, and \( z \in \text{int}(K) \cap H \). Assume further that \( \text{Vol}(K^{z} \cap H^{+}) = \text{Vol}(K^{z} \cap H^{-}) \) (we say \( H \) is medial for \( K \)). Then

\[
\text{Vol}(S_{u}K)^{z} \geq \text{Vol}(K^{z}).
\]

**Proof.** For the proof we shall invoke the Harmonic-mean Prékopa-Leindler type inequality for functions on \( \mathbb{R}^{+} \), which we proved in Section 3.5. Let us quote it again we only need \( \lambda = 1/2 \)

Let \( f, g, h : \mathbb{R}^{+} \rightarrow \mathbb{R}^{+} \) be integrable functions such that for every pair \((r, s) \in \mathbb{R}^{+} \times \mathbb{R}^{+} \)

\[
h\left(\frac{2rs}{s + r}\right) \geq f(r)\frac{s}{s + r}g(s)\frac{r}{s + r}.
\]

Then,

\[
\int_{0}^{\infty} h \geq \frac{2}{\int_{0}^{\infty} f \int_{0}^{\infty} g}.
\]

We assume without loss of generality that \( u = e_{n} \) and that \( z = 0 \). We shall next compare level sets in a similar way to the proof of Theorem 4.3, however we shall compare the levels \( K^{o}(r) \) and \( K^{o}(-r') \). We claim that for \( r, r' \geq 0 \) we have

\[
\frac{r'}{r + r'}K^{o}(r) + \frac{r}{r + r'}K^{o}(-r') \subseteq (S_{u}K)^{o}\left(\frac{2rr'}{r + r'}\right).
\]

Once this is established, we may define \( h(r) = \text{Vol}_{n-1}((S_{u}K)^{o}(r)) \), \( f(r) = \text{Vol}_{n-1}(K^{o}(r)) \) and \( g(r) = \text{Vol}_{n-1}(K^{o}(-r)) \). By Brunn-Minkowski inequality applied to the left hand side of the inclusion we get that these three functions satisfy

\[
h\left(\frac{2rr'}{r + r'}\right) \geq f(r)\frac{r'}{r + r'}g(r')\frac{r}{r + r'},
\]

which is precisely the condition in the Harmonic Prékopa Leindler type theorem we quoted above. By the theorem we get that

\[
\text{Vol}((S_{u}K)^{o}) = 2\int h(r) \geq \frac{2}{\int f \int g} \int f \int g
\]

However, since we assumed \( H \) satisfies that the total mass of \( K^{o} \cap H^{+} \) is half the total mass, we get that on the right hand side is (twice) the harmonic average of two equal numbers, so that in the end result

\[
\text{Vol}((S_{u}K)^{o}) \geq \text{Vol}(K^{o}).
\]
4.5 Completing the proof for the non-symmetric case

We shall first need one more proposition which is very similar to Proposition 4.10, but restricted to a subspace.

**Proposition 4.12.** Given some affine subspace $E$ with $\text{int}(K) \cap E \neq \emptyset$, the Function $F(z) = \text{Vol}(K^z) : \text{int}(K) \cap E \to \mathbb{R}^+$ (so, restricted to an affine subspace $E$) has a unique minimum. (This point $z_0(E)$ is called the Santaló point of $K$ with respect to $E$). It is attained on the unique point $z_0 \in E$ for which

$$\int_{K^{z_0} - z_0} x dx \in E^\perp.$$

**Proof.** The proof is identical to the proof of Proposition 4.10 above. Clearly $F$ is still a strictly convex function so it has a unique minimum. We have already computed the gradient of $F$ when defined on all of $\mathbb{R}^n$, and for a minimum under a constraint the gradient must belong to the subspace $E^\perp$. 

We shall use the above proposition to show following simple Lemma:

**Lemma 4.13.** If a convex body $K$ has its Santaló point with respect to a subspace $E$ at $z_0 \in E \cap K$, and if $u \in E^\perp$, and if $L$ is defined by $L^{z_0} = S_u(K^{z_0})$, then $z_0$ is the Santaló point of $L$ with respect to $E$.

**Proof.** All we need to check, by Proposition 4.12, is that $\int_{L^{z_0} - z_0} x dx \in E^\perp$. For simplicity we may assume $z_0 = 0$, $u = e_n$ and $E = (e_j, \ldots, e_n)^\perp$. So, we know that 0 is the Santalo point of $K$ with respect to $E$, so that $\int_{K^0} x dx \in E^\perp$, and we need to show that the same holds for $L$, where $L^0 = S_u K^0$. The latter implies that the $e_j$-component of the integral, $a_i = \int_{K^0} x_i dx = \int_{L^0} x_i dx$ for every $i \neq n$ (since Steiner symmetrization with respect to $e_n^\perp$ only changes the $n^{th}$ co-ordinate of each point) and in particular we have by assumption $a_i = 0$ for $i = 1, \ldots, j - 1$, as needed. 

We are now given a convex body $K$ which is compact and with non-empty interior. We shall apply $n$ orthogonal Steiner symmetrization, so get a centrally symetric body, and we will do this in such a way so as not to increase the volume product. To this end, let us first choose some unit vector basis $(u_i)_{i=1}^n$. We first move choose $b_1$ in such a way that $H_1 = u_1^\perp + b_1 u_1$ intersects $S$ so that half its volume lies in $H^+$ and half in $H^-$. (The Steiner symmetrization with respect to $H_1$ we denote by $S_{H_1}$, since $S_{u_1}$ would imply the subspace is through 0 and we want it to be an affine subspace through $x_1 u_1$, we could instead “move $K$”, but this way it is more elegant). We denote by $z_1 \in K \cap H_1$ the unique point for which $\text{Vol}(K^{z_1})$ is minimized. We shall denote by $K_1$ the body such that

$$S_{H_1}(K^{z_1}) = K_1^{z_1}.$$

That is, $K_1$ is obtained by taking the Steiner symmetrization of $K^{z_1}$ and taking its $z_1$-polar. Clearly $\text{Vol}(K_1^{z_1}) = \text{Vol}(K^{z_1})$, since $S_{H_1}$ is volume preserving. Also, since $H_1$ was medial for $K$ and $z_1 \in H$ we have by Lemma 4.11 that

$$\text{Vol}(K) = \text{Vol}((S_{H_1}(K^{z_1}))^{z_1}) \geq \text{Vol}((K^{z_1})^{z_1}) = \text{Vol}(K).$$

We got a new body, $K_1$, which is symmetric with respect to $H_1$, and for which

$$\text{Vol}(K) \text{Vol}(K^{z_1}) \leq \text{Vol}(K_1) \text{Vol}(K_1^{z_1}).$$
We shall inductively define $z_i \in \bigcap_{j=1}^{i} H_j$ where $H_j = u_j^+ + b_j u_j$ and a body $K_i$ which is symmetric with respect to $H_j, j = 1, \ldots, i$ so that

$$\text{Vol}(K_i) \text{Vol}(K_i^{z_i}) \leq \text{Vol}(K_{i+1}) \text{Vol}(K_{i+1}^{z_{i+1}}).$$

If we succeed, then at step $n$ we will get a body which is centrally symmetric about the point $z_n = \bigcap_{j=1}^{n} H_j$ and so $\text{Vol}(K^{z_n}) \text{Vol}(K) \leq \kappa_n^2$, by the centrally symmetric Santalo inequality which we have already proved.

For the induction step we choose the affine hyperplane $H_i = u_i^+ + b_i u_i$ which is orthogonal to $u_i$ and medial for $K_{i-1}$. We shall define $z_i$ to be the Santalo point for $K_{i-1}$ with respect to the subspace $\bigcap_{j=1}^{i} H_j$, and $K_i^{z_i} = S_{u_i}(K_{i-1}^{z_{i-1}})$. To show that $\text{Vol}(K_i) \text{Vol}(K_i^{z_i})$ is an increasing sequence we note first that

$$\text{Vol}(K_{i+1}^{z_{i+1}}) = \text{Vol}(S_{u_{i+1}}(K_{i+1}^{z_{i+1}})) = \text{Vol}(K_{i+1}^{z_{i+1}}) = \inf\{\text{Vol}(K_i^{z_i}) : z \in \bigcap_{j=1}^{i+1} H_j\}$$

$$\geq \inf\{\text{Vol}(K_i^{z_i}) : z \in \bigcap_{j=1}^{i+1} H_j\} = \text{Vol}(K_{i+1}^{z_{i+1}}).$$

(This step was equality in step 1, since we did not define $z_0$ to be the absolute best point for $K$ in all of $\mathbb{R}^n$, which we could, in principle). By Lemma 4.11, since $H_{i+1}$ is medial for $K_i = (K_{i+1}^{z_{i+1}})^{z_{i+1}}$ we have that

$$\text{Vol}((S_{H_{i+1}}(K_{i+1}^{z_{i+1}}))^{z_{i+1}} \geq \text{Vol}((K_{i+1}^{z_{i+1}}))^{z_{i+1}} = \text{Vol}(K_{i+1}).$$

But the body on the left hand side is our definition of $K_{i+1}$. Putting the two together we get

$$\text{Vol}(K_i) \text{Vol}(K_i^{z_i}) \leq \text{Vol}(K_{i+1}) \text{Vol}(K_{i+1}^{z_{i+1}}).$$

□

If the proof seems a bit long and tedious, all that is really going on is $n$ Steiner symmetrization, chosen correctly, in a way where the “centers with which one polarizes” are kept in record. The feeling of complication might be remedied when we give the functional version and proof, which are similar and yet, in a sense, cleaner.

4.6 Lower bounds

4.6.1 In the plane

I remind you, Mahler’s conjecture is that

$$s(K) = \inf_{x} \{\text{Vol}_n(K) \text{Vol}_n((K - x)^0)\}$$

is minimized for $K$ which is a simplex, and among centrally symmetric convex bodies, for $K$ which is a cube (but other bodies give the same value as the cube, for example $B_1^n$, but others too). For a simplex the value is $(n + 1)(n + 1)/(n!)$, and for the cube $4^n/n!$. Both these conjectures are open for all $n > 2$. Some things are known though. Reisner showed the inequality (centrally symmetric) for zonotopes (zonoids), these are bodies which are (limits of) unit segments (in particular they are centrally symmetric). Also, for unit balls of 1-unconditional norms (in other words, they have many assumed symmetries: w.r.t. all coordinate hyperplanes). For the ones which have the symmetry group of the simplex, the non-symmetric Mahler is known. It is also known that for symmetric, the cube is a local minimizer, and in general, the simplex...
is (locally). Finally, the minimizer cannot have a point with positive Gauss curvature (never mind now the definitions).

Despite the course’s intentions to be mainly directed toward high dimensional phenomena, I am tempted to give you the two dimensional proof of Mahler’s inequality, for the chance that one of you might have a good idea how to attack the three dimensional question.

**Proposition 4.14.** For every triangle $0 \in \Delta \subset \mathbb{R}^2$ we have $s(\Delta) \geq 27/4$ with equality iff the center of mass is at 0. Also, for every $\alpha > 27/4$ there exists a triangle with $s(\Delta) = \alpha$.

**Proof.** Let the triangle have vertices $x, y, z$ and without loss of generality (by affine invariance) assume $x = e_1$ and $y = e_2$ and let $z = (-s, -t)^T$ with $s, t > 0$.

The dual body is thus a triangle with one edge on $x = 1$ one on $y = 1$ and the last on $-sx - ty = 1$. So, one vertex is $(1, 1)$ and the others (compute) are $(1, -(1 + s)/t)$ and $(-(1 + t)/s, 1)$.

Its area is easily computed to be equal $(1 + s + t)^2/2st$. Thus

$$s(\Delta) = (1 + s + t)^3/4st.$$  

Use the arithmetic geometric means inequality to check that $(1 + s + t)/3 \geq (st)^{1/3}$ and thus $(1 + s + t)^3/4st \geq 3^3/4 = 27/4$.

For equality we clearly need $s = t = 1$ in which case the center of mass is at 0. For the last assertion note that as $t \to \infty$ and $s$ is fixed (say, $s = 1$) we get the ratio approaching $+\infty$. $\square$

The proof of the lower bound for general polytopes will follow by reduction of the number of vertices. The key ingredient is

**Proposition 4.15.** Let $\ell_1, \ell_2$ be two half-lines with common endpoint $w$ and including an angle $\theta < \pi$ at $w$. Let $z$ be some point between the two rays, and let $x \in \ell_1$ and $y \in \ell_2$ such that $z \in [x, y]$. The area of the triangle $x, y, w$ is minimal when $z$ is the center of $[x, y]$. If $x$ is allowed to vary only on some interval in $\ell_1$ (above $au$ otherwise there will be no corresponding $y$), the maximal volume is attained at an endpoint of the interval. Similarly for $y$.

**Proof.** This is a simple extremization problem. Assume $w = 0$ and the rays in directions $u$ and $v$. Let $z = \alpha u + \beta v$ for $\alpha, \beta > 0$, and by abuse of notation let $x = xu$ and $y = yv$ then the condition is that $z = (1 - \lambda)xu + \lambda yv$ for some $\lambda \in [0, 1]$. Thus $x = 1/(1 - \lambda)\alpha$ and $y = 1/(\lambda\beta)$. The area of the triangle is $xy\sin \theta$, and thus we maximize their product, which means $\lambda = 1/2$. $\square$

**Corollary 4.16** (Mahler’s proof). For a polytope $0 \in P \subset \mathbb{R}^2$ with $n \geq 4$ vertices there is a polytope $P'$ with $(n - 1)$ vertices such that $s(P') < s(P)$.

**Proof.** Let the vertices of $P$ be $x_1, x_2, \ldots, x_n$ ($x_{n+1} := x_1$), the dual polygon $P^\circ$ has edges corresponding to these vertices, given by $\langle \cdot, x_i \rangle = 1$ and vertices which we shall call $\xi_i$ so that $[\xi_i, \xi_{i+1}]$ is on $\langle \cdot, x_i \rangle = 1$. Consider the angles from 0, and call $\phi_i$ the angle $x_i, 0, x_{i+1}$. If $n \geq 5$ then there is at least some $i$ with $\phi_i + \phi_{i+1} < \pi$. If $n = 4$ then there might not be, but this is only if these angles are all $\pi/2$, which means that $P^\circ$ is a parallelogram (draw). If 0 is in its center, then we know $v(P) = 8 > 27/4$. If 0 is not in its center, exchange $P$ and $P^\circ$ and find two consecutive angles with sum less than $\pi$.

So, assume we have these two angles. Assume these are $\phi_1$ and $\phi_2$. That the angles sum up to less than $\pi$ implies that the dual lines $\langle \cdot, x_1 \rangle = 1$ and $\langle \cdot, x_3 \rangle = 1$ are not parallel. Consider the segment $[x_1, x_3]$ and the line through $x_2$ which is parallel to this segment. It meets the two
lines prolonged from the edges $[x_n, x_1]$ and $[x_3, x_4]$ at some points $y_1, y_2$ such that $x_2 \in [y_1, y_2]$. For every $\lambda \in [0, 1]$ we define a new polygone with the same vertices except for $x_2$ which will become $(1 - \lambda)y_1 + \lambda y_2$. All these polygones have the same area. The dual polygones $(P(t))^\circ$ now have edges on the same lines as before except for the edge corresponding to $x_2$, which still passes through the point corresponding to the line through $[y_1, y_2]$ and joining $\langle \cdot, x_1 \rangle = 1$ and $\langle \cdot, x_3 \rangle = 1$. The total area is thus the area up to the intersection point (at some $\xi_0$, say) minus the area of a triangle with new endpoints $\xi_1(\lambda)$ and $\xi_2(\lambda)$. These are allowed to travel on some fixed interval on the lines $\langle \cdot, x_1 \rangle = 1$ and $\langle \cdot, x_3 \rangle = 1$ so that the point corresponding to the line through $[y_1, y_2]$ remains on the segment between them. By the previous proposition, this is attained at an endpoint, namely $\lambda = 0$ or $\lambda = 1$. \hfill \Box

In Exercise sheet 5 you are asked to prove in this method the centrally symmetric Mahler in $\mathbb{R}^2$. 
4.6.2 General dimension

As mentioned above, the question of the optimal lower bound for \( s(K) \), both in the centrally symmetric case, and in the general case, is open for dimension above 2. Let us discuss the family of conjectured minimizers for the centrally symmetric case.

The first two members in this family are the cube \([-1, 1]^n\), and its dual, the cross-polytope \( B_1^n = \text{conv}(\pm e_i, i = 1, \ldots, n) \) (along with all of their linear images, of course). Next, given two symmetric bodies \( K \subset \mathbb{R}^k \) and \( T \subset \mathbb{R}^m \) which are conjectured minimizers for the \( k \)- and \( m \)-dimensional questions (that is, give values \( 2^k/k! \) and \( 2^m/m! \) respectively), one constructs a body in dimension \( n = m + k \) as follows:

\[
L = K \times T = \{(x, y) \in \mathbb{R}^k \times \mathbb{R}^m : x \in K, y \in T\}.
\]

We claim that \( L \), and its dual of course, would then be (conjectured) minimizers in \( \mathbb{R}^n \). In other words, they will give the same volume product as the cube. To this end let us understand their structure and volume.

**Lemma 4.17.** Let \( K \subset \mathbb{R}^k \) and \( T \subset \mathbb{R}^m \), then

\[
(K \times T)^o = \text{conv}(K^o \times \{0\}, \{0\} \times T^o).
\]

**Proof.** We simply write

\[
(K \times T)^o = \{(x, y) : \langle x, w \rangle + \langle y, z \rangle \leq 1 \forall w \in K, z \in T\}
\]

\[
= \bigcup_{0 \leq \lambda \leq 1} \{(x, y) : \langle x, w \rangle \leq 1 - \lambda \text{ and } \langle y, z \rangle \leq \lambda \forall w \in K, z \in T\}
\]

\[
= \bigcup_{0 \leq \lambda \leq 1} \{(x, y) : x \in (1 - \lambda)K^c \text{ and } y \in \lambda T^c\}
\]

\[
= \bigcup_{0 \leq \lambda \leq 1} \{((1 - \lambda)x, 0) + \lambda(0, y) : x \in K^c \text{ and } y \in T^c\}
\]

as needed. \( \square \)

The volume of \( L = K \times T \) is \( \text{Vol}_n(L) = \text{Vol}_k(K)\text{Vol}_m(T) \). The volume of the convex hull is also easily computed as follows.

**Lemma 4.18.** Let \( K \subset \mathbb{R}^k \) and \( T \subset \mathbb{R}^m \), then

\[
\text{Vol}_{m+k}(\text{conv}(K \times \{0\}, \{0\} \times T) = \frac{1}{\binom{n}{k}} \text{Vol}_k(K)\text{Vol}_m(T).
\]

**Proof.** To compute this volume we simply integrate, say with respect to the \( y \) coordinate, the volume of the fiber above \( y \). That is, we should decide what is the \( k \)-dimensional volume of the set

\[
\{x : (x, y) \in \text{conv}(K \times \{0\}, \{0\} \times T)\}.
\]

We check which is the minimal \( \lambda \) such that \( y = \lambda y' \) for \( y' \in T \), this is clearly simply \( \|y\|_T \), and then the fiber consists of a copy of \( (1 - \lambda)K \). Therefore, the total volume is

\[
\int_T (1 - \|y\|_T)^k \text{Vol}(K)dy.
\]

This integral can be written in polar coordinates.
\[
\int_T (1 - \|y\|_T)^k dy = m\kappa_m \int_{S^{m-1}} \int_0^\infty 1_T(\rho u)(1 - r\|u\|_T)^k r^{m-1} drd\sigma(u)
\]

\[
= m\kappa_m \int_{S^{m-1}} \int_0^{1/\|u\|_T} (1 - r\|u\|_T)^k r^{m-1} drd\sigma(u)
\]

\[
= m\kappa_m \int_{S^{m-1}} \|u\|^{-m} d\sigma(u) \int_0^1 (1 - s)^k s^{m-1} ds
\]

\[
= m\Vol(T) \int_0^1 (1 - s)^k s^{m-1} ds = \Vol(T) mB(m + 1)
\]

\[
= \Vol(T)m\frac{(m-1)!k!}{m+k!} = \Vol(T)\frac{1}{\binom{m+k}{m}}.
\]

Combining with the computation above, we our stated value. \qed

**Corollary 4.19.** Let \( K \subset \mathbb{R}^k \) and \( T \subset \mathbb{R}^m \) be such that

\[
\Vol(K)\Vol(K^\circ) = 2^k/k! \quad \text{and} \quad \Vol(T)\Vol(T^\circ) = 2^m/m!,
\]

then for \( L = K \times T \subset \mathbb{R}^n \) (with \( n = m + k \)) we have \( \Vol(L)\Vol(L^\circ) = 2^n/n! \).

**Proof.**

\[
\Vol(L)\Vol(L^\circ) = \Vol(K)\Vol(T)\Vol(K^\circ)\Vol(T^\circ)\frac{1}{\binom{n}{k}} = \frac{2^k 2^m m!k!}{k! m! n!} = \frac{2^n}{n!}.
\]

\qed

The Hanner polytopes are the polytopes acheived by this procedure, and its dual, starting with \( K \) and \( T \) which are segments in \( \mathbb{R} \). They are conjectured to be the minimizers, in the world of centrally symmetric convex sets, of the volume product.

We mention that by a similar trick one may show that a “weak” lower bound, anything better than exponential, would imply the correct lower bound. See the exercise sheet.

### 4.7 Functional Blaschke-Santalo inequality

The classical Blaschke-Santalo inequality, which was described above, states that the volume product \( s(K) \) is maximaized for Euclidean balls and their affine images. Euclidean balls (of radius 1) are the only self-dual bodies in \( \mathbb{R}^n \).

For log-concave functions, with duality given by the Legendre duality discussed in Section 3.4, one thus anticipates an inequality of similar form. Define for an even function the quantity

\[
s(\varphi) = \int \exp^{-\varphi} \int \exp^{-L\varphi}.
\]

Note that when considering an upper bound we may as well restrict to convex functions, since \( \varphi^{**} \leq \varphi \) and thus \( s(\varphi) \leq s(\varphi^*) \). As you checked in Exercice 4, the only self-dual function on \( \mathbb{R}^n \) is \( \varphi(x) = -|x|^2/2 \), for which we get \( s(\varphi) = \left( \int \exp(-|x|^2/2) \right)^2 = (2\pi)^n \). One may similarly define for general functions the expression

\[
s(\varphi) = \inf_{x \in \mathbb{R}^n} s(\varphi(\cdot - x)),
\]

34
and it turns out that these expressions maximize when the function $\varphi$ is self-dual. We have

**Theorem 4.20.** Let $\varphi : \mathbb{R}^n \to (-\infty, +\infty]$ be a proper function, and assume that
\[
\int x \exp(-\varphi(x)) dx = 0 \text{ then }
\int \exp(-\varphi) \int \exp(-L\varphi) \leq (2\pi)^n.
\]

### 4.7.1 Proof in the even case

The even case follows from an integration argument, using the standard Prékopa Leindler inequality. We mention to begin with that one may assume that $\varphi$ is a convex lower semi continuous function, by replacing $\varphi$ by $L\varphi$ and using that $LL\varphi \leq \varphi$.

We next make a reduction to non-negative functions, since replacing $\varphi$ by $\varphi + C$ changes $L\varphi$ to $L\varphi - C$ and thus the integrals are multiplied by constants $e^C$ and $e^{-C}$ so the volume product remains intact. Therefore, we may assume $\varphi(0) = 0$, and in such a case this is its minimal value since $\varphi$ is both even and convex.

We use that $\varphi(x) + L\varphi(y) \geq \langle x, y \rangle$ which implies that if $L\varphi(y) \leq s$ then for every $x$ with $\varphi(x) \leq t$ we have $\langle x, y \rangle \leq s + t$ so that
\[
\left\{ y : L\varphi(y) \leq s \right\} \subset (s + t)\left\{ x : \varphi(x) \leq t \right\}.
\]

Using the Blaschke-Santalo inequality we see that
\[
\text{Vol}\left(\left\{ y : L\varphi(y) \leq s \right\}\right)\text{Vol}\left(\left\{ x : \varphi(x) \leq t \right\}\right) \leq (s + t)^n \kappa^n
\]
where we used that the functions are even and thus these sets are centrally symmetric and no translation is needed.

Note that
\[
\int e^{-\varphi} = \int_1^{\infty} \text{Vol}\left\{ x : \exp(-\varphi(x)) \geq r \right\} dr = \int_1^{\infty} \text{Vol}\left\{ x : \varphi(x) \leq \ln \frac{1}{r} \right\} dr = \int_0^{\infty} \text{Vol}\left\{ x : \varphi(x) \leq t \right\} e^{-t} dt.
\]
Here we have used our assumption that $\varphi$ is minimized at 0 where it assumes the value 0.

We thus set $f(t) = e^{-t}\text{Vol}\left\{ x : \varphi(x) \leq t \right\}$, $g(s) = e^{-s}\text{Vol}\left\{ y : L\varphi(y) \leq s \right\}$, here $f, g : \mathbb{R}^+ \to \mathbb{R}^+$, and our inequality states that for $s, t > 0$ we have
\[
f(t)g(s) \leq e^{-(s+t)}(s + t)^n \kappa^n.
\]

Letting $h(r) = e^{-r}(2r)^{n/2}\kappa^n : \mathbb{R}^+ \to \mathbb{R}^+$ we get that
\[
h\left(\frac{s + t}{2}\right) \geq (f(t)g(s))^{1/2},
\]
and by the Prékopa Leindler inequality,
\[
\int h \geq \left( \int f \int g \right)^{1/2}
\]

35
which translates to
\[
\int e^{-\varphi} \int e^{-L\varphi} \leq \left( \int h \right)^2 = \left( \kappa \int_0^\infty e^{-r} (2r)^{n/2} dr \right)^2 = (2\pi)^n.
\]

Here we have used either an integral of a special function, or the realization that for \( \varphi(x) = |x|^2/2 \) we have total integral \((2\pi)^2\) and sub level sets of \( \varphi \) are \( \{ x : |x|^2 < 2r \} \) with volume \( \kappa_n(2r)^{n/2} \) hence integral against \( \exp(-r) \) is exactly \((2\pi)^n\).

### 4.7.2 General case

As usual, the general case is slightly more involved. This time we do not work with Steiner symmetrizations, and do not reduce to the symmetric case (though this is in fact possible, and the original proof was in fact structured that way). The theorem will follow from the slightly more general

**Theorem 4.21.** Let \( f \) and \( g \) be non-negative Borel functions on \( \mathbb{R}^n \) satisfying
\[
\forall x, y \in \mathbb{R}^n, \quad f(x)g(y) \leq e^{-\langle x, y \rangle},
\]
and assume that \( \int xg(x) = 0 \). Then
\[
\int_{\mathbb{R}^n} f(x) dx \int_{\mathbb{R}^n} g(y) dy \leq (2\pi)^n.
\]

**Remark 4.22.** Note first that by simply integrating the inequality on \( f \) and \( g \) with respect to \( x \) and \( y \) we get nothing, since the integral of \( e^{-\langle x, y \rangle} \) is infinite. Next, note that for \( n = 1 \) if \( f(s)g(t) \leq \exp(-st) \) on \( \mathbb{R}^+ \), this means \( F(x) = f(e^x)e^x \) and \( G(y) = g(e^y)e^y \) satisfy on \( \mathbb{R} \) that
\[
F(x)G(y) = f(e^x)g(e^y)e^{x+y} \leq e^{-\exp(x+y)} e^{x+y} = H^2(\frac{x+y}{2}),
\]
where \( H(z) = e^{-\frac{1}{4} z^2} e^z \), and so by the Prékopa-Leindler inequality we get
\[
\int_0^\infty f \int_0^\infty g = \int_{-\infty}^\infty F \int_{-\infty}^\infty G \leq \left( \int_{-\infty}^\infty H \right)^2 = \int_0^\infty e^{-r^2/2} dr = 2\pi.
\]

To prove Theorem 4.21 we shall use a proposition which has the falour of Lemma 4.11. We prove the following

**Theorem 4.23.** Let \( f \) be a non-negative Borel function on \( \mathbb{R}^n \) with \( 0 < \int f < \infty \) and \( xf(x) \) integrable as well. Let \( H \) be an affine hyperplane splitting \( \mathbb{R}^n \) into two half-spaces \( H_+ \) and \( H_- \). Define \( \lambda \in [0, 1] \) by \( \lambda \int_{\mathbb{R}^n} f = \int_{H_+} f \). Then there exists \( z \in H \) such that: for every non-negative Borel function \( h \) such that \( f(z + x)h(y) \leq e^{-\langle x, y \rangle} \) for all \( x, y \in \mathbb{R}^n \), one has
\[
\int_{\mathbb{R}^n} f \int_{\mathbb{R}^n} h \leq \frac{1}{4\lambda(1 - \lambda)} (2\pi)^n.
\]

Let us first explain why Theorem 4.23 implies Theorem 4.21. We only need the case \( \lambda = 1/2 \). We are given \( f \) and \( g \) with \( \int xg(x) = 0 \), such that \( f(x)g(y) \leq \exp(-\langle x, y \rangle) \). The fact that \( f \) is integrable as is \( xf(x) \) follows from the fact that the barycenter of \( g \) is at the origin. Indeed, there are points \( (y_i)_{i=1}^m \) for which \( g(y_i) \neq 0 \) with \( 0 \in \text{conv}(y_i) \) and therefore letting \( c = \min g(y_i) \)
we have \( f(x) \leq \frac{1}{c} \min \exp(-\langle x, y_i \rangle) \) and this latter function had a converging integral, and also when multiplied by \(|x|\).

We choose some hyperplane \( H \) which splits \( \int f \) into two equal pieces (if the total integral is 0 there is nothing to prove). Find the \( z \) from the conclusion of the theorem.

Letting \( h(y) = g(y)\exp(\langle y, z \rangle) \) we have

\[
 f(z + x)h(y) = f(z + x)g(y)\exp(\langle y, z \rangle) \leq \exp(-\langle z + x, y \rangle)\exp(\langle y, z \rangle) = \exp(-\langle x, y \rangle).
\]

Thus we may apply the conclusion of Theorem 4.23 to \( f \) and \( h \) and get that

\[
 \int_{\mathbb{R}^n} f(x)dx \int_{\mathbb{R}^n} g(y)\exp(\langle y, z \rangle)dy \leq (2\pi)^n.
\]

Finally, using that the barycenter of \( g \) is at the origin, we have that \( \int_{\mathbb{R}^n} g(y)\exp(\langle y, z \rangle)dy \) is minimal for \( z = 0 \), which can be done either by derivation (very similar to the case of bodies) or directly, since

\[
 \int g(y)dy \leq \int g(y)(\exp(\langle y, z \rangle) - \langle y, z \rangle)dy = \int g(y)\exp(\langle y, z \rangle)dy.
\]

Thus, the required inequality is satisfied.

To prove Theorem 4.23 we shall use following multiplicative version of the Prékopa Leindler inequality (we actually used it above in the one dimensional case without explicitly naming it)

**Theorem 4.24** (Prékopa Leindler Multiplicative). Let \( f, g, h : \mathbb{R}^+ \to \mathbb{R} \) and assume that \( f^{1/2}(s)g^{1/2}(t) \leq h(st)^{1/2} \). Then

\[
 \left( \int f \right)^{1/2} \left( \int g \right)^{1/2} \leq \int h.
\]

**Proof.** We let \( F(x) = f(e^x)e^x \) and \( G(y) = g(e^y)e^y \) and \( H(z) = h(e^z)e^z \) be defined on \( \mathbb{R} \) so that these satisfy

\[
 F(x)G(y) = f(e^x)g(e^y)e^{x+y} \leq h^2(e^{x+y})e^{x+y} = H \left( \frac{x+y}{2} \right)^2.
\]

We see that \( F, G, H \) satisfy the conditions of the usual Prékopa Leindler inequality, thus

\[
 \int_{\mathbb{R}} F \int_{\mathbb{R}} G \leq (\int H)^2
\]

which may be rewritten as

\[
 \int_{\mathbb{R}^+} f \int_{\mathbb{R}^+} g \leq (\int h)^2.
\]

**Corollary 4.25.** Let \( f, g : \mathbb{R}^+ \to \mathbb{R} \) and assume that \( f(s)g(t) \leq \exp(-st) \). Then

\[
 \int f \int g \leq \frac{\pi}{2}.
\]

**Proof of Theorem 4.23.** The proof goes by induction on dimension. The case of dimension 1 follows from the multiplicative Prékopa Leindler theorem. Indeed, assume first \( \int f = 1 \). For any fixed \( r \in \mathbb{R} \) we let \( \lambda = \int_{r}^{\infty} f \). We shall pick \( z = r \). Indeed, assume that \( g \) satisfies
f(r + s)g(t) \leq \exp(-st) \text{ for all } s, t. \text{ Then by applying the multiplicative version to } f(r + s) \text{ and } g(t), \text{ we get } \int_{-\infty}^{\infty} f f_{1}^{\infty} g \leq \pi/2, \text{ and by applying it to } f(r - s) \text{ and } g(-t) \text{ we get that } \int_{-\infty}^{r} f f_{0}^{\infty} g \leq \pi/2. \text{ Using the definition of } \lambda \text{ we see that }

\int_{\mathbb{R}} g \leq \frac{\pi}{2\lambda} + \frac{\pi}{2(1 - \lambda)} = \frac{2\pi}{4\lambda(1 - \lambda)}.

In higher dimensions, we are given some \( H \) and \( \lambda \) is chosen so then \( \lambda \int_{\mathbb{R}^n} f = \int_{H} f. \) Assume \( \lambda \neq 0 \) and \( \lambda \neq 1 \) (in which case the inequality is useless anyway).

We consider the barycenters of \( f \) in \( H^+, \) say \( b^+ \), and in \( H^- \), say \( b^- \). The line between then intersects \( H \) at a unique point, which will be our chosen \( z \). Without loss of generality \( z = 0 \), since we may translate \( f \) and \( H \). So, assume that for some \( g \) we have \( f(x)g(y) \leq \exp(-\langle x, y \rangle) \) for all \( x, y \) (This \( g \) was called \( h \) in the original statement, but since the hyperplane is also \( H \), we switched names).

For simplicity denote \( H = e_n^+ \), and assume \( \langle e_n, b^+ \rangle > 0 \). Let \( v = b^+/\langle e_n, b^+ \rangle \), that is, a vector of “height” 1 in direction \( b^+ \). Let \( A \) be the linear operator sending \( e_n \) to \( v \) and not changing \( e_j \) for \( j = 1, \ldots, n \). Denote \( B = (A^{-1})^* \). Check that \( Be_n = e_n \).

Define lower dimensional functions on \( H \) itself, using integration, as follows

\[
F_+(y) = \int_{R^+} f(y + sv)dy \quad G_+(y') = \int_{R^+} g(By' + te_n)dt.
\]

This is done so that the barycenter of \( F_+ \) is at the origin (this is why we integrate “along \( v \)”) and so that the arguments will satisfy

\[
\langle y + sv, By' + te_n \rangle = \langle A(y + se_n), B(y' + te_n) \rangle = \langle y + se_n, y' + te_n \rangle = \langle y, y' \rangle + ts
\]

Thus by our assumption

\[
f(y + sv)g(By' + te_n) \leq \exp(-\langle y, y' \rangle - st)
\]

so we may apply the one dimensional \( \mathbb{R}^+ \) case to get that for any \( y, y' \in H \) we have

\[
F_+ G_+(y') \leq \frac{\pi}{2} \exp(-\langle y, y' \rangle).
\]

Now, using that the barycenter of \( F_+ \) is 0, we have by the induction assumption (of the whole theorem) that

\[
\int_{H} F_+ \int_{H} G_+ \leq \frac{\pi}{2} (2\pi)^{n-1}.
\]

By Fubini, since \( \det(A) = 1 \), we have \( \int_{H} F_+ = \int_{H} f = \lambda \int f \) so this implies that (using that \( \det(B) = 1 \))

\[
\int_{H} g = \int_{H^+} g(Bx)dx \leq \frac{1}{4\lambda} (2\pi)^n.
\]

We do the same for \( H^- \), so that

\[
\int g \leq \left( \frac{1}{4\lambda} + \frac{1}{4(1 - \lambda)} \right) (2\pi)^n = \frac{(2\pi)^n}{4\lambda(1 - \lambda)}.
\]

\( \square \)
We mention yet another approach to the Blaschke-Santalo inequality for functions, which is function-symmetrization. for lack of time, let me only explain it in brief. A log-concave function can be approximated by $1/s$-concave functions with $s \to \infty$. Indeed, if $f(x_0) = \exp(-\varphi(x))$ for a convex function $\varphi$, let

$$f_s(x) = (1 - \frac{\varphi(x)}{s})^{s+1}.$$ 

Here $A_s = \max(A, 0)$. Then clearly $f_s(x) \to \exp(-\varphi(x))$ as $s \to \infty$, locally uniformly, and clearly $f_s^{1/s}$ is concave.

For an $s$-concave function on $\mathbb{R}^n$, we can consider an associated body in $\mathbb{R}^{n+s}$ which is convex, and such that its measure projection (a.k.a. marginal) onto $\mathbb{R}^n$ gives $f_s$. There are many such bodies, but only one which is round in the remaining $s$ directions. It is given by

$$K_s(f_s) = \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^s : x \in \text{supp}(f), |y| \leq \kappa_s^{-1/s}f_s^{1/s}(x)\}.$$ 

One may apply usual polarity transform to $K_s(f_s)$ and obtain a new body, $K_s^\circ(f_s)$, which is also round in directions of $\mathbb{R}^s$, and corresponds to another function

$$\pi_{\mathbb{R}^n} K_s^\circ(f_s)(x) = \text{Vol}_n (K_s^\circ(f_s) \cap (x + \mathbb{R}^s)) = g_s(x).$$ 

We thus get a “duality-transform” for $1/s$-concave functions (we know that the marginal of a uniform measure on a convex body onto $s$-codimensions is $1/s$-concave).

By the usual Santalo inequality, the volume product for these functions is maximal when they are marginals of the uniform volume on dual ellipsoids. Then a limiting argument should be added. I gave some parts of this (without the limiting arguments) in exercise sheet 6.

5 Back to Brunn-Minkowski

We start with a couple of applications of Brunn-Minkowski inequality, and then we shall see a couple more proofs of it.

5.1 Rogers Shephard inequality

**Definition 5.1.** The difference body of a convex body $K$ is the centrally symmetric convex body

$$K - K = \{x - y : x, y \in K\}.$$ 

From the Brunn-Minkowski inequality it is clear that $\text{Vol}_n(K - K) \geq 2^n \text{Vol}_n(K)$, with equality if and only if $K$ has a centre of symmetry. Rogers and Shephard gave a sharp upper bound for the volume of the difference body.

**Theorem 5.2** (Rogers-Shephard). Let $K$ be a convex body in $\mathbb{R}^n$. Then,

$$\text{Vol}_n(K - K) \leq \left(\frac{2^n}{n}\right) \text{Vol}_n(K).$$

**Proof.** The Brunn-Minkowski inequality enters the proof through the observation that $f(x) = \text{Vol}_n(K \cap (x + K))^{1/n}$ is a concave function supported on $K - K$. This can be proved directly, using the inclusion

$$K \cap ((1 - \lambda)x + \lambda y + T) \supset (1 - \lambda)(K \cap (x + T)) + \lambda(K \cap (y + T)).$$

39
and the Brunn-Minkowski inequality, or by using Brunn’s concavity principle for the body \( K \times T \subset \mathbb{R}^{2n} \) and the \( n \)-dimensional subspace \( \{(x, x) : x \in \mathbb{R}^n\} \) (in both cases one employs the special case \( T = K \)).

We define a second function \( g : K - K \to \mathbb{R}^+ \) as follows: each \( x \in K - K \) can be written in the form \( x = r\theta \), where \( \theta \in S^{n-1} \) and \( 0 \leq r \leq \rho_{K-K}(\theta) \). Recall that \( \rho_W \) denotes the radial function of \( W \):

\[
\rho_W(\theta) = \max\{t > 0 : t\theta \in W\}, \quad \theta \in S^{n-1}.
\]

We then set \( g(x) = f(0)(1 - r/\rho_{K-K}(\theta)) \). By definition, \( g \) is linear on the interval \([0, \rho_{K-K}(\theta)\theta]\), it vanishes on the boundary of \( K - K \), and \( g(0) = f(0) \). Since \( f \) is concave, we see that \( f \geq g \) on \( K - K \). Therefore, we can write

\[
\int_{K-K} \text{Vol}_n(K \cap (x + K))dx = \int_{K-K} f^n(x)dx \geq \int_{K-K} g^n(x)dx
\]

\[
= \left[f(0)\right]^n n\kappa_n \int_{S^{n-1}} \int_0^{\rho_{K-K}(\theta)} r^{n-1} (1 - r/\rho_{K-K}(\theta))^n dr d\sigma(\theta)
\]

\[
= n\kappa_n \text{Vol}_n(K) \int_{S^{n-1}} \rho_{K-K}(\theta)^n d\sigma(\theta) \int_0^1 t^{n-1}(1 - t)^n dt
\]

\[
= \text{Vol}_n(K) \text{Vol}_n(K - K) \frac{\Gamma(n)\Gamma(n+1)}{\Gamma(2n+1)}
\]

\[
= \left(\frac{2n}{n}\right)^{-1} \text{Vol}_n(K) \text{Vol}_n(K - K),
\]

using integration in polar coordinates (we denote by \( \sigma \) the rotationally invariant probability measure on the Euclidean unit sphere \( S^{n-1} \), and use the fact that its surface area is \( \text{Vol}_{n-1}(S^{n-1}) = n\kappa_n \)). On the other hand, Fubini’s theorem gives

\[
\int_{K-K} \text{Vol}_n(K \cap (x + K))dx = \int_{\mathbb{R}^n} \text{Vol}_n(K \cap (x + K))dx
\]

\[
= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} 1_K(y)1_{x+K}(y)dy dx = \int_{\mathbb{R}^n} 1_K(y) \left(\int_{\mathbb{R}^n} 1_{y-K}(x)dx\right) dy
\]

\[
= \int_K \text{Vol}_n(y - K)dy = \text{Vol}_n(K)^2.
\]

Combining the above, we conclude the proof. \( \square \)

**Remark 5.3.** If we take a closer look at the argument and take into account the equality case in the Brunn-Minkowski inequality, we see that we can have equality in Theorem 5.2 if and only if \( K \) has the following property: if \((rK+x) \cap (sK+y) \neq \emptyset\) for some \( r, s > 0 \) and \( x, y \in \mathbb{R}^n \), then

\[
(rK+x) \cap (sK+y) = tK + w
\]

for some \( t \geq 0 \) and \( w \in \mathbb{R}^n \). That is, the non-empty intersection of homothetical copies of \( K \) is again homothetical to \( K \) or a point. Rogers and Shephard proved that this property characterizes the simplex.

The usefulness of Theorem 5.2 rests upon the fact that the volume of \( K - K \) is not much larger than the volume of \( K \). We have

\[
\text{Vol}_n(K - K)^{1/n} \leq 4\text{Vol}_n(K)^{1/n},
\]

which means that every convex body (that contains the origin) is contained in a centrally symmetric convex body with more or less the same volume radius; the volume radius of a convex body is the radius of the Euclidean ball centered at 0 that has the same volume, or, in other words, it is the quantity \( (\text{Vol}_n(K)/\kappa_n)^{1/n} \). You will show something similar in the reverse direction in the exercise sheet.
It is good to note that Rogers and Shephard also proved that, when the barycenter of \( K \) is 0, we additionally have that

\[
\text{Vol}_n(K \cap (-K)) \geq 2^{-n}\text{Vol}_n(K).
\]

So, every convex body with barycenter at the origin is not only contained in a centrally symmetric convex body of similar volume radius, but also contains a centrally symmetric convex body with the same volume radius up to a constant. Again this result has a generalization as follows:

\[
\text{Vol}_n(K) \text{Vol}_n(L) \leq \text{Vol}_n(K + L) \text{Vol}_n(K \cap (-L)).
\]

One more, connected, inequality of Rogers and Shephard, which we next state and prove.

**Lemma 5.4.** Let \( K \) be a centrally symmetric convex body in \( \mathbb{R}^n \). Let \( F \in G_{n,k} \) and write \( F^\perp \) for the orthogonal complement of \( F \). Then,

\[
\left( \binom{n}{k} \right)^{-1} \text{Vol}_k(K \cap F) \text{Vol}_{n-k}(P_{F^\perp}(K)) \leq \text{Vol}_n(K) \leq \text{Vol}_k(K \cap F) \text{Vol}_{n-k}(P_{F^\perp}(K)).
\]

**Proof.** Using Fubini’s theorem we write

\[
(13) \quad \text{Vol}_n(K) = \int_{P_{F^\perp}(K)} \text{Vol}_k(K \cap (x + F)) \, dx.
\]

From the symmetry of \( K \) we have

\[
\text{Vol}_k(K \cap (x + F)) = \text{Vol}_k(K \cap (-x + F))
\]

for every \( x \in P_{F^\perp}(K) \). Since

\[
\frac{1}{2}K \cap (x + F) + \frac{1}{2}K \cap (-x + F) \subseteq K \cap F,
\]

by the Brunn-Minkowski inequality it follows that

\[
\text{Vol}_k(K \cap (x + F)) \leq \text{Vol}_k(K \cap F)
\]

for all \( x \in P_{F^\perp}(K) \), which proves the right hand side inequality in (13).

For the left hand side inequality we first observe, using the convexity of \( K \), that if \( x \in tP_{F^\perp}(K) \) for some \( 0 \leq t \leq 1 \) then there exists \( z \in F \) such that \( x + tz \in tK \), and hence

\[
K \cap (x + F) \supseteq (1 - t)(K \cap F) + x + tz.
\]

Taking volumes, we get

\[
\text{Vol}_k(K \cap (x + F)) \geq (1 - \|x\|_{F,K})^k \text{Vol}_k(K \cap F)
\]

for every \( x \in P_{F^\perp}(K) \), where \( \| \cdot \|_{F,K} \) is the norm induced by \( P_{F^\perp}(K) \) on \( F^\perp \). Thus, the right hand side of (13) is greater than or equal to

\[
\text{Vol}_k(K \cap F) \text{Vol}_{n-k}(P_{F^\perp}(K)) \int_0^1 k t^{n-k}(1 - t)^{k-1} \, dt.
\]
Computing the last integral we complete the proof of the lemma.

5.2 Borell’s lemma

Borell’s lemma implies concentration of volume in convex bodies in $\mathbb{R}^n$: if $A \cap K$ captures more than half of the volume of $K$, then the percentage of $K$ that stays outside $tA$, when $t > 1$, decreases exponentially with respect to $t$ as $t \to \infty$, at a rate that does not depend at all on the body $K$ or the dimension $n$.

**Theorem 5.5** (Borell’s Lemma). Let $K$ be a convex body in $\mathbb{R}^n$ with volume $\text{Vol}_n(K) = 1$, and let $A$ be a closed, convex and centrally symmetric set such that $\text{Vol}_n(K \cap A) = \delta > \frac{1}{2}$. Then, for every $t > 1$ we have

$$\text{Vol}_n(K \cap (tA)^c) \leq \delta \left( \frac{1 - \delta}{\delta} \right)^{\frac{t+1}{2}}.$$

The same is true when $\text{Vol}(A \cap K)$ is replaced by any log-concave probability measure $\mu$.

**Proof.** We first show that

$$A^c \supseteq \frac{2}{t+1} (tA)^c + \frac{t-1}{t+1} A.$$

If this were not so, we could write some $a \in A$ in the form

$$a = \frac{2}{t+1} y + \frac{t-1}{t+1} a_1,$$

for some $a_1 \in A$ and $y \notin tA$. But then we would have

$$\frac{1}{t} y = \frac{t+1}{2t} a + \frac{t-1}{2t} (-a_1) \in A,$$

because of the convexity and symmetry of $A$. This means that $y \in tA$, which is a contradiction.

Since $K$ is convex, we have

$$A^c \cap K \supseteq \frac{2}{t+1} ((tA)^c \cap K) + \frac{t-1}{t+1} (A \cap K).$$

An application of the Brunn-Minkowski inequality gives

$$1 - \delta = \text{Vol}_n(A^c \cap K) \geq \text{Vol}_n(((tA)^c \cap K)^{\frac{2}{t+1}} \text{Vol}_n(A \cap K)^{\frac{t-1}{t+1}}$$

$$= \text{Vol}_n((tA)^c \cap K)^{\frac{2}{t+1}} \delta^{\frac{t-1}{t+1}}.$$

This proves the theorem. \(\square\)

You will show in the next exercise sheet that there is no magic to the number $1/2$ appearing in the theorem, and a similar statement will hold when it is replaces by some other constant, say $1/3$.

Since this lemma here is somewhat unattached to the other topics, let me show an example of how it may be used to prove something. This “thing” is called a Khinchine-type inequality, and it says basically that if you want to average some norm on a convex body (or a log-concave measure) then it does not matter which $L_p$ norm you take, up to a factor $cq/p$. It is a sort of reverse-Hölder inequality.
Proposition 5.6. Let $f$ be a norm on $\mathbb{R}^n$ and let $\mu$ be a log concave probability measure. Let $1 \leq p \leq q$. Then
\[
\left( \int f^p d\mu \right)^{1/p} \leq \left( \int f^q d\mu \right)^{1/q} \leq c \frac{q}{p} \left( \int f^p d\mu \right)^{1/p},
\]
where $c$ is a universal constant independent on $f, p, q, n$.

Proof. We shall use the set
\[
A = \{ x : f(x) \leq 3 \left( \int f^p d\mu \right)^{1/p} \}.
\]
Clearly $tA$ is the set where $f(x) \leq 3t \left( \int f^p d\mu \right)^{1/p}$. The measure of $A$ is bounded using Markov by
\[
1 - \mu A = \mu \{ x : f^p(x) \geq 3 \left( \int f^p d\mu \right)^{1/p} \} \leq 3^{-p} \leq 1/3.
\]
By Borell’s lemma $1 - \mu(tA) \leq e^{-cpt}$ and thus
\[
\int_{\mathbb{R}^n} f^q d\mu = \int_0^\infty q s^{q-1} \mu(f > s) ds.
\]
Denoting $3 \left( \int f^p d\mu \right)^{1/p} = A_p$ we write
\[
\int_0^\infty q s^{q-1} \mu(f > s) ds \leq \int_0^{A_p} q s^{q-1} ds + \int_{A_p}^\infty q s^{q-1} e^{-csp/A_p} ds
\]
and continue to get the estimate. \qed

5.3 Measure transportation proofs

Let $K$ and $T$ be two open convex bodies in $\mathbb{R}^n$. By a volume preserving transformation we mean a map $\phi : K \to T$ which is one to one, onto and has a Jacobian with constant determinant equal to $\text{Vol}_n(T)/\text{Vol}_n(K)$. In this section we describe two such maps, the Knothe map and the Brenier map. Applying each one of them we may obtain alternative proofs of the Brunn-Minkowski inequality.

5.3.1 Knothe map

We fix an orthonormal basis $\{e_1, \ldots, e_n\}$ in $\mathbb{R}^n$, and consider two open convex bodies $K$ and $T$. The properties of the Knothe map from $K$ to $T$ with respect to the given coordinate system are described in the following theorem.

Theorem 5.7. There exists a map $\phi : K \to T$ with the following properties:

(a) $\phi$ is triangular: the $i$-th coordinate function of $\phi$ depends only on $x_1, \ldots, x_i$. That is,
\[
\phi(x_1, \ldots, x_n) = (\phi_1(x_1), \phi_2(x_1, x_2), \ldots, \phi_n(x_1, \ldots, x_n)).
\]

(b) The partial derivatives $\frac{\partial \phi_i}{\partial x_i}$ are positive on $K$, and the determinant of the Jacobian matrix
$J(\phi)$ of $\phi$ is constant. More precisely, for every $x \in K$

$$|\det J(\phi)(x)| = \prod_{i=1}^{n} \frac{\partial \phi_i}{\partial x_i}(x) = \frac{\text{Vol}_n(T)}{\text{Vol}_n(K)}.$$  

**Proof.** For each $i = 1, \ldots, n$ and $s = (s_1, \ldots, s_i) \in \mathbb{R}^i$ we consider the section

$$K_s = \{y \in \mathbb{R}^{n-i} : (s,y) \in K\}$$

of $K$ (similarly for $T$). We shall define a one to one and onto map $\phi : K \to T$ as follows. Let $x = (x_1, \ldots, x_n) \in K$. Then, $K_{x_1} \neq \emptyset$ and we can define $\phi_1(x) = \phi_1(x_1)$ by

$$\frac{1}{\text{Vol}_n(K)} \int_{-\infty}^{x_1} \text{Vol}_{n-1}(K_{s_1})ds_1 = \frac{1}{\text{Vol}_n(T)} \int_{-\infty}^{\phi_1(x_1)} \text{Vol}_{n-1}(T_{t_1})dt_1.$$  

In other words, we move in the direction of $e_1$ until we “catch” a percentage of $T$ which is equal to the percentage of $K$ occupied by $K \cap \{s = (s_1, \ldots, s_n) : s_1 \leq x_1\}$. Note that $\phi_1$ is defined on $K$ but $\phi_1(x)$ depends only on the first coordinate of $x \in K$. Also,

$$\frac{\partial \phi_1}{\partial x_1}(x) = \frac{\text{Vol}_n(T)}{\text{Vol}_n(K)} \frac{\text{Vol}_{n-1}(K_{x_1})}{\text{Vol}_{n-1}(T_{\phi_1(x_1)})}.$$  

We continue by induction. Assume that we have defined $\phi_1(x) = \phi_1(x_1), \phi_2(x) = \phi_2(x_1, x_2)$ and $\phi_{j-1}(x) = \phi_{j-1}(x_1, \ldots, x_{j-1})$ for some $j \geq 2$. If $x = (x_1, \ldots, x_n) \in K$ then $K(x_1, \ldots, x_{j-1}) \neq \emptyset$, and we define $\phi_j(x) = \phi_j(x_1, \ldots, x_j)$ by

$$\frac{\text{Vol}_{n-j+1}(T_{(\phi_1(x_1), \ldots, \phi_{j-1}(x_1, \ldots, x_{j-1}))})}{\text{Vol}_{n-j+1}(K(x_1, \ldots, x_{j-1}))} \int_{-\infty}^{x_j} \text{Vol}_{n-j}(K(x_1, \ldots, x_{j-1}, s_j))ds_j$$

$$= \int_{-\infty}^{\phi_j(x_1, \ldots, x_j)} \text{Vol}_{n-j}(T_{(\phi_1(x_1), \ldots, \phi_{j-1}(x_1, \ldots, x_{j-1}), t_j)})dt_j.$$  

It is clear that

$$\frac{\partial \phi_j}{\partial x_j}(x) = \frac{\text{Vol}_{n-j+1}(T_{(\phi_1(x), \ldots, \phi_{j-1}(x))})}{\text{Vol}_{n-j+1}(K(x_1, \ldots, x_{j-1}))} \frac{\text{Vol}_{n-j}(K(x_1, \ldots, x_j))}{\text{Vol}_{n-j}(T_{(\phi_1(x), \ldots, \phi_j(x))})}.$$  

Continuing in this way, we obtain a map $\phi = (\phi_1, \ldots, \phi_n) : K \to T$. It is easy to check that $\phi$ is one to one and onto. Note that

$$\frac{\partial \phi_n}{\partial x_n}(x) = \frac{\text{Vol}_1(T_{(\phi_1(x), \ldots, \phi_{n-1}(x))})}{\text{Vol}_1(K(x_1, \ldots, x_{n-1}))}.$$  

By construction, $\phi$ has properties (a) and (b).  

**Remark.** Observe that each choice of coordinate system in $\mathbb{R}^n$ produces a different Knothe map from $K$ onto $T$.

Using the Knothe map we can give one more proof of the Brunn-Minkowski inequality for convex bodies. We may clearly assume that $K$ and $T$ are open. Consider the Knothe map
φ : K → T. It is clear that

\[(\text{Id} + \phi)(K) \subseteq K + \phi(K) = K + T,\]

and, since J(Id + φ) is triangular, we get

\[
\text{Vol}_n(K + T) \geq \int_{\text{Id} + \phi(K)} dx = \int_K |\det J(\text{Id} + \phi)(x)| dx = \int_K \prod_{j=1}^n \left(1 + \frac{\partial \phi_i}{\partial x_j}(x)\right)^{1/n} dx
\]

\[
\geq \int_K \left(1 + \left(\prod_{j=1}^n \frac{\partial \phi_i}{\partial x_j}(x)\right)^{1/n}\right)^n dx = \text{Vol}_n(K) \left(1 + \left(\frac{\text{Vol}_n(T)}{\text{Vol}_n(K)}\right)^{1/n}\right)^n
\]

\[
= \left(\text{Vol}_n(K)^{1/n} + \text{Vol}_n(T)^{1/n}\right)^n.
\]

### 5.3.2 Brenier map

In this section we show yet another way to produce a mapping ψ : K → T which is volume preserving (by this I mean constant determinant of Jacobian, not necessarily 1 if Vol(K) ≠ Vol(T)). This time we will make sure that the Jacobian of ψ is positive definite, by ensuring ψ = ∇f for a (twice differentiable) convex function f. Using this map (called the Brenier map) we shall have one more proof of the Brunn-Minkowski inequality for convex bodies as follows: Since (Id + ψ)(K) ⊆ K + T,

\[
\text{Vol}_n(K + T) \geq \int_K |\det J(\text{Id} + \psi)(x)| dx = \int_K |\det (\text{Id} + \text{Hess}f)(x)| dx
\]

\[
= \int_K \prod_{i=1}^n (1 + \lambda_i(x)) dx,
\]

where λ_i(x) are the non negative eigenvalues of Hess f. Moreover, by the ratio-of-volumes-preserving property of ψ, we have \(\prod_{i=1}^n \lambda_i(x) = \text{Vol}_n(T)/\text{Vol}_n(K)\) for every \(x \in K\). Therefore, the arithmetic-geometric means inequality gives

\[
\text{Vol}_n(K + T) \geq \int_K \left(1 + \prod_{i=1}^n \lambda_i(x)^{1/n}\right)^n dx = (\text{Vol}_n(K)^{1/n} + \text{Vol}_n(T)^{1/n})^n.
\]

So, one must find a mapping which is of the form \(\psi = \nabla f\) for a convex function \(f\) defined on \(K\), the image of which is \(T\).

**Definition 5.8** (Rockafellar). Let \(G \subseteq \mathbb{R}^n \times \mathbb{R}^n\). We say that \(G\) is cyclically monotone if for every \(m \geq 2\) and \((x_i, y_i) \in G, i \leq m,\) we have

\[
\langle y_1, x_2 - x_1 \rangle + \langle y_2, x_3 - x_2 \rangle + \cdots + \langle y_m, x_1 - x_m \rangle \leq 0.
\]

We shall next link the property of cyclic monotonicity with that of being contained in the graph of the sub-gradient of some convex function. To this end we remind what is the definition of the sub-differential of a function:
**Definition 5.9.** Let \( \varphi : \mathbb{R}^n \to (-\infty, +\infty] \) be convex and \( x \in \text{dom}(\varphi) \). The sub-differential of \( \varphi \) at \( x \) is defined by

\[
\partial \varphi(x) = \{ u \in \mathbb{R}^n : \varphi(y) \geq \varphi(x) + \langle u, y-x \rangle \text{ for all } y \}.
\]

Any \( u \in \partial \varphi(x) \) is called a sub-gradient of \( \varphi \) at \( x \).

The subdifferential parametrizes the supporting hyperplanes of \( \varphi \). It is easy to see that \( \nabla \varphi(x) \) exists if and only if \( \partial \varphi(x) = \{ \nabla \varphi(x) \} \). Let \( x \in \text{int}(\text{dom}(\varphi)) \). Then the set \( \partial \varphi(x) \) is non-empty, since there is at least one supporting hyperplane at \( (x, \varphi(x)) \), this hyperplane being the graph of an affine function which lies below \( \varphi \) and touches it at \( (x, \varphi(x)) \). The set \( \partial \varphi(x) \) is convex. For \( x \in \text{int}(\text{dom}(\varphi)) \), it is compact (namely it is closed, and at the same time \( \partial \varphi(x) \) cannot include a ray since \( x \) is in the interior of the domain, thus it is bounded).

**Proposition 5.10** (Rockafellar). Let \( G \subseteq \mathbb{R}^n \times \mathbb{R}^n \). Then, \( G \) is contained in the graph of the sub-differential of a proper convex function \( \varphi : \mathbb{R}^n \to \mathbb{R} \) if and only if \( G \) is cyclically monotone.

**Proof.** It is easy to check that the subdifferential of a proper convex function is cyclically monotone. Let \( (x_i, y_i) \) be contained in the graph of \( \partial \varphi \), \( i = 1, \ldots, m \) (namely let \( y_i \in \partial \varphi(x_i) \)). Then,

\[
\langle y_1, x_2 - x_1 \rangle \leq \varphi(x_2) - \varphi(x_1)
\]
\[
\langle y_2, x_3 - x_2 \rangle \leq \varphi(x_3) - \varphi(x_2)
\]
\[
\vdots
\]
\[
\langle y_m, x_1 - x_m \rangle \leq \varphi(x_1) - \varphi(x_m)
\]

by the definition of the subdifferential. Adding the inequalities we get

\[
\langle y_1, x_2 - x_1 \rangle + \langle y_2, x_3 - x_2 \rangle + \cdots + \langle y_m, x_1 - x_m \rangle \leq 0.
\]

It follows that every \( G \) which is a subset of the graph of \( \partial \varphi \) is cyclically monotone.

For the opposite direction, assume that \( G \) is non-empty and cyclically monotone, and fix \( (x_0, y_0) \in G \). We define \( \varphi : \mathbb{R}^n \to \mathbb{R} \) by

\[
\varphi(x) = \sup \{ \langle y_m, x - x_m \rangle + \langle y_{m-1}, x_m - x_{m-1} \rangle + \cdots + \langle y_0, x_1 - x_0 \rangle \},
\]

where the supremum is taken over all \( m \geq 0 \) and \( (x_i, y_i) \in G, 1 \leq i \leq m \). The function \( \varphi \) is convex since it is the supremum of a family of affine functions. Using the cyclic monotonicity of \( G \) we easily check that \( \varphi(x_0) = 0 \). This shows that \( \varphi \) is proper. Finally, \( G \) is contained in the sub-differential of \( \varphi \): Let \( (x, y) \in G \). We will show that

\[
t + \langle z - x, y \rangle < \varphi(z)
\]

for every \( t < \varphi(x) \) and every \( z \in \mathbb{R}^n \). This implies that \( y \in \partial \varphi(x) \). Since \( t < \varphi(x) \), there exist \( (x_1, y_1), \ldots, (x_m, y_m) \in G \) such that

\[
t < \langle y_m, x - x_m \rangle + \cdots + \langle y_0, x_1 - x_0 \rangle.
\]
By the definition of $\varphi$ again,
\[
\varphi(z) \geq \langle y, z - x \rangle + \langle y_m, x - x_m \rangle + \cdots + \langle y_0, x_1 - x_0 \rangle > \langle y, z - x \rangle + t.
\]

This completes the proof. \qed

**Remark 5.11.** We mention that for a convex l.s.c. function $\varphi$ we have $L(\varphi) + \varphi(x) = \langle x, y \rangle$ if and only if $y \in \partial \varphi(x)$ if and only if $x \in \partial L(\varphi)$. Indeed, if $y \in \partial \varphi(x)$ then for all $z$ we have $\varphi(z) \geq \varphi(x) + \langle y, z - x \rangle$ which means $\sup \langle z, x \rangle - \varphi(z) = \langle x, y \rangle - \varphi(x)$ which means $L(\varphi) + \varphi(x) = \langle x, y \rangle - \varphi(x)$ as claimed. The other direction is similarly direct.

Recall that we are searching for a ratio-of-volumes-preserving mapping $\psi = \nabla \varphi$. Let us first generalize the notion of “volume preserving” to mappings that push forward one measure into another. The special case of uniform measures on convex sets will be of particular interest.

**Definition 5.12.** Consider the space $\mathcal{P}(\mathbb{R}^n)$ of Borel probability measures on $\mathbb{R}^n$ as a subset of the unit ball of $C^\infty(\mathbb{R}^n)^*$ (the dual of the space of infinitely differentiable functions which vanish uniformly at infinity). Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$. We say that a probability measure $\gamma \in \mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n)$ has marginals $\mu$ and $\nu$ if for every bounded Borel measurable functions $f, g : \mathbb{R}^n \to \mathbb{R}$ we have
\[
\int_{\mathbb{R}^n} f(x) d\mu(x) = \int_{\mathbb{R}^n \times \mathbb{R}^n} f(x) d\gamma(x, y)
\]
and
\[
\int_{\mathbb{R}^n} g(y) d\nu(y) = \int_{\mathbb{R}^n \times \mathbb{R}^n} g(y) d\gamma(x, y).
\]

**Proposition 5.13.** Let $\mu$ and $\nu$ be Borel probability measures on $\mathbb{R}^n$. There exists a joint probability measure $\gamma$ on $\mathbb{R}^n \times \mathbb{R}^n$ which has cyclically monotone support and marginals $\mu, \nu$.

The proof of Proposition 5.13 is a consequence of the next discrete lemma and a limiting argument.

**Lemma 5.14.** Let $x_i, y_i \in \mathbb{R}^n$, $i = 1, \ldots, m$ and consider the measures
\[
\mu = \frac{1}{m} \sum_{i=1}^{m} \delta_{x_i} \quad \text{and} \quad \nu = \frac{1}{m} \sum_{i=1}^{m} \delta_{y_i}.
\]

There exists a probability measure $\gamma$ on $\mathbb{R}^n \times \mathbb{R}^n$ which has cyclically monotone support and marginals $\mu, \nu$.

**Proof.** For every permutation $\sigma$ of $\{1, \ldots, m\}$ we consider the measure
\[
\gamma_\sigma = \frac{1}{m} \sum_{i=1}^{m} \delta_{(x_{\sigma(i)}, y_i)}.
\]

It is clear that $\gamma_\sigma$ has marginals $\mu$ and $\nu$ for every $\sigma$. Let
\[
F(\sigma) = \sum_{i=1}^{m} \langle y_i, x_{\sigma(i)} \rangle.
\]
We will show that if \( F(\sigma) \) is maximal among all permutations, then the support \((x_{\sigma(i)}, y_i)\) of \( \gamma_\sigma \) is cyclically monotone.

Without loss of generality we may assume that \( F(I) \) is maximal, where \( I \) denotes the identity permutation. We want to show that \( G = \{(x_i, y_i) : i \leq m\} \) is cyclically monotone. Let \( k \leq m, i_1, \ldots, i_k \) be distinct indices and consider the points \((x_{i_1}, y_{i_1}) \in G\). If \( \sigma \) is the permutation defined by \( \sigma(i_s) = i_{s+1} \) if \( s < k \), \( \sigma(i_k) = i_1 \) and \( \sigma(i) = i \) otherwise, we have

\[
0 \geq F(\sigma) - F(I) = \sum_{s=1}^{k} (\langle y_{i_s}, x_{\sigma(i_s)} \rangle - \langle y_{i_s}, x_{i_s} \rangle) \\
= \langle y_{i_1}, x_{i_2} - x_{i_1} \rangle + \langle y_{i_2}, x_{i_3} - x_{i_2} \rangle + \cdots + \langle y_{i_k}, x_{i_1} - x_{i_k} \rangle.
\]

This proves the lemma.  

**Proof of Proposition 5.13.** Given \( \mu \) and \( \nu \) we construct discrete measures \( \mu_n \to \mu, \nu_n \to \nu \) which converge in the weak-* topology. By Lemma 5.14, for every \( n \) there exists \( \gamma_n \) with cyclically monotone support and marginals \( \mu_n, \nu_n \). By a standard compactness argument (Arzela-Ascoli) there exists a weak-* subsequential limit \( \gamma \) of \( \gamma_n \). It clearly has cyclically monotone support and \( \mu, \nu \) as its marginals.

**Definition 5.15.** Let \( \mu, \nu \in \mathcal{P}(\mathbb{R}^n) \). Let \( T : \mathbb{R}^n \to \mathbb{R}^n \) be a measurable function which is defined \( \mu \)-almost everywhere and satisfies

\[
\nu(B) = \mu(T^{-1}(B))
\]

for every Borel subset \( B \) of \( \mathbb{R}^n \). We then say that \( T \) pushes forward \( \mu \) to \( \nu \) and write \( T\mu = \nu \).

It is easy to see that \( T\mu = \nu \) if and only if for every bounded Borel measurable \( g : \mathbb{R}^n \to \mathbb{R} \) we have

\[
\int_{\mathbb{R}^n} g(y)d\nu(y) = \int_{\mathbb{R}^n} g(T(x))d\mu(x).
\]

Given \( \mu \) and \( \nu \in \mathcal{P}(\mathbb{R}^n) \), we have the next theorem establishing the existence of a map that is the gradient of a convex function and pushes forward \( \mu \) to \( \nu \). Brenier proved its existence and uniqueness under some integrability assumptions on the moments of \( \mu \) and \( \nu \); these were later removed by McCann. The precise formulation is as follows.

**Theorem 5.16** (Brenier-McCann). Let \( \mu, \nu \in \mathcal{P}(\mathbb{R}^n) \) and assume that \( \mu \) is absolutely continuous with respect to the Lebesgue measure. Then, there exists a convex function \( \varphi : \mathbb{R}^n \to \mathbb{R} \) such that \( \nabla \varphi : \mathbb{R}^n \to \mathbb{R}^n \) is defined \( \mu \)-almost everywhere, and \( \nabla \varphi \mu = \nu \).

**Proof.** Proposition 5.13 shows that there exists a probability measure \( \gamma \) on \( \mathbb{R}^n \times \mathbb{R}^n \) which has cyclically monotone support and marginals \( \mu, \nu \). By Proposition 5.10, the support of \( \gamma \) is contained in the subdifferential of a proper convex function \( \varphi : \mathbb{R}^n \to \mathbb{R} \).

Since \( \varphi \) is convex and \( \mu \) is absolutely continuous with respect to the Lebesgue measure, \( \varphi \) is differentiable \( \mu \)-almost everywhere. Since \( \text{supp}(\gamma) \subset \partial \varphi \), by the definition of the subdifferential we have \( y = \nabla \varphi(x) \) for almost all pairs \((x, y)\) with respect to \( \gamma \). Then, for every bounded Borel measurable \( g : \mathbb{R}^n \to \mathbb{R} \) we see that

\[
\int g(y)d\nu(y) = \int g(y)d\gamma(x, y) = \int g(\nabla \varphi(x))d\gamma(x, y) = \int g(\nabla \varphi(x))d\mu(x),
\]

48
which shows that \((\nabla \varphi)\mu = \nu\).

\[ \nabla \varphi \mu = \nu. \]

**Remark 5.17.** Regularity results from the theory of elliptic equations provide smoothness of the Brenier map. In the particular case that \(\mu\) is the normalized Lebesgue measure on some convex body \(K\) and \(\nu\) is the normalized Lebesgue measure on some other convex body \(T\), a regularity result of Caffarelli shows that \(\varphi\) may be assumed twice continuously differentiable. More precisely, we have the following theorem, from which one can deduce the Brunn-Minkowski inequality as explained at the beginning of this subsection.

Let \(K\) and \(T\) be open convex bodies in \(\mathbb{R}^n\). There is a convex function \(\varphi \in C^2(K)\) such that \(\psi = \nabla \varphi : K \rightarrow T\) is one to one, onto and volume preserving.

An interesting variant of this theorem was given in a paper of Alesker Dar and V. Milman. They built a volume preserving map between convex bodies which has the following remarkable property.

Let \(K\) and \(T\) be open convex bounded subsets of \(\mathbb{R}^n\) of volume 1. Then, there exists a \(C^1\)-diffeomorphism \(F : K \rightarrow T\) preserving the Lebesgue measure, such that, for every \(\lambda > 0\), \(\{x + \lambda F(x) : x \in K\} = K + \lambda T\).

As an illustrative example for the meaning of this statement, consider two \(\varepsilon\)-extensions of two orthogonal intervals in \(\mathbb{R}^2\). Then the map \(x \mapsto x + F(x)\) tends to a Peano type curve as \(\varepsilon \rightarrow 0\). The function \(F\) cannot be in general the Brenier map, as the example of \(K\) being a ball and \(T\) an ellipsoid shows. Indeed, in such a case the Brenier map is linear but then \(\text{Id} + F\) is also linear, whereas the sum of a ball and an ellipsoid is not, in general, an ellipsoid. However, \(F\) in this theorem is the composition of two Brenier maps, mapping the uniform measure on \(K\) to the Gaussian measure (say) and mapping the Gaussian measure to the uniform measure on \(T\).

**5.3.3 Another proof sketch for Brenier’s theorem**

A beautiful proof using Brower’s fixed point theorem for Brenier’s theorem exists, and I would like to describe it to you, non-rigorously. What’s nice about this proof is that it also illustrates the relation to the cost function in the transportation. We start with an example in \(\mathbb{R}^2\): Imagine you want to transport the uniform measure on a convex body to two point masses (whose sum equals the total mass of the body). Imagine you want to do this minimizing the square cost

\[ \int_A |x - p_0|^2 dx + \int_B |x - p_1|^2 dx \]

over all partitions \(K = A \cup B\). We claim that the minimizer is attained by dividing \(K\) into two parts using a hyperplane orthogonal to \(e = p_1 - p_0\).

Indeed, the quadratic cost somehow “separates” the coordinates, and with any other scheme you could exchange two pieces with different \(e\) component and decrease to total cost.

The transportation scheme we came up with is, turns out, the gradient of a (piecewise linear, with just two pieces) convex function. More generally, whenever we find a transport map which is optimal with respect to the quadratic cost, it will be the gradient of a convex function: this is cyclic monotonicity exactly! For any permutation \(\sigma \in S_m\) we may rearrange
\( x_i \) so that \( \sigma \) send \( i \) to \( i + 1 \) and \( m \) to \( 1 \) and then

\[
\sum_{i=1}^{m} |x_i - p_i|^2 - \sum_{i=1}^{m} |x_{\sigma(i)} - p_i|^2 = \langle p_1, x_2 - x_1 \rangle + \langle p_2, x_3 - x_2 \rangle + \cdots + \langle p_m, x_1 - x_m \rangle \leq 0.
\]

So, cyclic monotonicity implies optimality with respect to quadratic distances.

**Theorem 5.18.** If \( \mu \) and \( \nu \) are probability measures on \( \mathbb{R}^n \), \( \nu \) has compact support and \( \mu \) assigns no mass to any set of Hausdorff dimension \((n - 1)\), then there is a convex function \( \varphi: \mathbb{R}^n \to \mathbb{R} \), so that \( T = \nabla \varphi \) transports \( \mu \) to \( \nu \).

The proof has two main steps. We first establish the existence of \( \varphi \) in the case that the second measure \( \nu \) is discrete and then pass to general measures by approximating them weakly by discrete ones. Suppose then that \( \nu \) is a convex combination of point masses \( \nu = \sum_{i=1}^{m} \lambda_i \delta_{p_i} \). The convex function we want has the form

\[
\varphi(x) = \max_{1 \leq i \leq m} (\langle x, p_i \rangle - s_i).
\]

for some \( s_1, s_2, \ldots, s_m \in \mathbb{R} \). This function partitions \( \mathbb{R}^n \) into \( m \) pieces according to whichever linear function is biggest. If \( A_i \) is the set where the maximum is attained for index \( i \), then we want to juggle the \( s_i \) so as to arrange that \( \mu(A_i) = \lambda_i \) for each \( i \). In order to use a fixed point theorem we shall define a map \( H \) on the simplex of points \( t = (t_1, t_2, \ldots, t_m) \) with non-negative coordinates satisfying \( \sum t_i = 1 \), by considering the function

\[
\varphi_t(x) = \max \{ \langle x, p_i \rangle - \frac{1}{t_i} \}.
\]

If \( t \) has non-zero coordinates, \( H(t) \) will be the point \((\mu(A_1), \ldots, \mu(A_m))\) whose coordinates are the measures of the sets on which \( \varphi_t \) is linear. As \( t_i \to 0 \) clearly \( \mu(A_i) \) decreases to zero. In view of the hypothesis on \( \mu \), \( H \) can be defined continuously on the simplex and maps each face of the simplex into itself. It is a well-known consequence (or reformulation) of Brouwer's fixed point theorem, that such a map is surjective. (If the map omits a point, say in the interior of the simplex, then we can follow it by a projection of the punctured simplex onto its boundary, to obtain a continuous map from the simplex to its boundary which fixes each face. Now by cycling the coordinates of the simplex we obtain a continuous map with no fixed point.) So there is a choice of \( t \) for which \( H(t) \) is \((\lambda_1, \ldots, \lambda_m)\).

Now suppose that we have a general probability \( \nu \). Approximate it weakly by a sequence \((\nu_k)\) of discrete measures and choose convex functions \( \varphi_k \) whose gradients transport \( \mu \) to these approximants. Assume that the \( \varphi_k \) are pinned down by (for example) \( \varphi_k(0) = 0 \). We may assume that all \( \nu_k \) are supported on the same compact set from which it is immediate that the \( \varphi_k \) are equicontinuous. So we may assume that they have a locally uniform limit, \( \varphi \), say. The function \( \varphi \) is convex. A standard result in convex analysis guarantees that outside a set of Hausdorff dimension \( n \), \( \varphi \) and all the \( \varphi_k \) are differentiable: (for the \( \varphi_k \) the differentiability is obvious by their construction). It is quite easy to check that except on the exceptional set, \( \nabla \varphi_k \to \nabla \varphi \) Let \( T_k = \nabla \varphi_k \) for each \( k \) and \( T = \nabla \varphi \). By the condition on the support of \( \varphi \), \( T_k \to T \), \( \mu \)-almost everywhere. We want to conclude that \( T \) transports \( \mu \) to \( \nu \). This is a standard argument in weak convergence.
6 Classical positions of convex bodies

We fix a Euclidean structure \( \langle \cdot, \cdot \rangle \) in \( \mathbb{R}^n \). Given a convex body \( K \) in \( \mathbb{R}^n \), we consider the family of its positions, meaning its affine images \( \{ x_0 + T(K) \} \) where \( x_0 \in \mathbb{R}^n \) and \( T \in GL_n \). In the context of functional analysis, one is given a norm (whose unit ball is \( K \)) and the choice of a position reflects a choice of a Euclidean structure for the linear space \( \mathbb{R}^n \). In this case, when the body is centrally symmetric, one usually disregards translations and works only with linear images of \( K \). In this language, the choice of a Euclidean structure specifies a unit ball of the Euclidean norm, which is an ellipsoid. Thus, one may regard a “position” as a choice of a special ellipsoid. An ellipsoid in \( \mathbb{R}^n \) is a convex body of the form

\[
E = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^{n} \frac{(x, v_i)^2}{\alpha_i^2} \leq 1 \right\},
\]

where \( \{v_i\}_{i \leq n} \) is an orthonormal basis of \( \mathbb{R}^n \) with respect to \( \langle \cdot, \cdot \rangle \), and \( \alpha_1, \ldots, \alpha_n \) are positive reals (the directions and lengths of the semiaxes of \( E \), respectively). It is easy to check that \( E = T(B_2^n) \), where \( T \) is the positive definite linear transformation of \( \mathbb{R}^n \) defined by \( T(v_i) = \alpha_i v_i \), \( i = 1, \ldots, n \). Therefore, the volume of \( E \) is equal to

\[
\text{Vol}_n(E) = \kappa_n \prod_{i=1}^{n} \alpha_i
\]

(where, as before, \( \kappa_n = \text{Vol}_n(B_2^n) \)). Note that by standard linear algebra, for any \( A \in GL_n \), the body \( A(B_2^n) \) is an ellipsoid, and its volume is \( \kappa_n |\det(A)| \), and these are all ellipsoids.

Since we are discussing ellipsoids, and shall use them in the sequel, let us take a short detour regarding them. We shall sometimes think of a symmetric matrix as an element in \( \mathbb{R}^{n(n+1)/2} \). The subset of \( \mathbb{R}^{n(n+1)/2} \) corresponding to positive definite symmetric matrices is a convex cone (open, since we do not allow semi-definite). This cone has a very interesting extremal structure, its extremal rays (if one takes the closure, for convenience) are the rank one matrices, \( v^T v = v \otimes v \). It turns out that real symmetric matrices and the cone of positive definite matrices are a good model for Brunn-Minkowski type inequalities. Let us list a few examples.

The polynomiality of determinant is easy, simply by the formula for determinant. Polarization follows easily, we shall see an similar statement for volume in the sequel (I hope)

**Lemma 6.1.** Let \( A_1, \ldots, A_m \in GL_n(\mathbb{R}) \) be symmetric. Then

\[
\det(\lambda_1 A_1 + \cdots + \lambda_m A_m) = \sum_{i_1, \ldots, i_n=1}^{m} D(A_{i_1}, \ldots, A_{i_n}) \lambda_{i_1} \cdots \lambda_{i_n}.
\]

Here the mixed discriminant are

\[
D(A_1, \ldots, A_n) = \frac{1}{n!} \sum_{\varepsilon \in \{0,1\}^n} (-1)^n + \sum \varepsilon_i \det(\sum \varepsilon_i A_i).
\]

Note that in this lemma one does not assume positive matrices, and so one cannot claim that the coefficients, the mixed discriminants, are positive. For positive definite \( A_i \) we do have

\[
D(A_1, \ldots A_n) \geq \prod \det(A_i)^{1/n}.
\]
(We shall not prove, or use). We will use, however,

**Lemma 6.2.** Let $A_1, A_2 \in GL_n$ be symmetric and positive definite. Then

$$\det(A_1 + A_2)^{1/n} \geq \det(A_1)^{1/n} + \det(A_2)^{1/n}.$$  

Equality holds if and only if $A_1 = \mu A_2$ for some $\mu > 0$. Equivalently,

$$\det((1 - \lambda)A_1 + \lambda A_2) \geq \det(A_1)^{1-\lambda} \det(A_2)^{\lambda}$$

for all $\lambda \in [0,1]$. *Here the equality condition is that $A_1 = A_2$.***

This inequality can be proved as a consequence of the arithmetic-geometric means inequality in the same way as in the proof of Brunn Minkowski inequality for cubes. One only has to notice the fact from linear algebra that two positive definite linear transformations may be brought to diagonal form simultaneously by an $SL_n$ transform. We give here a different proof using H"older’s inequality.

**Proof.** We shall use the well known formula

$$\int_{R^n} e^{-\langle Ax, x \rangle} dx = \frac{(2\pi)^{n/2}}{\det(A)^{1/2}}.$$  

Then we write

$$\frac{(2\pi)^{n/2}}{\det((1 - \lambda)A + \lambda B)^{1/2}} = \int_{R^n} e^{-\langle((1-\lambda)A + \lambda B)x, x \rangle} dx$$

$$= \int_{R^n} \left(e^{-\langle Ax, x \rangle}(1-\lambda)\left(e^{-\langle Bx, x \rangle}\right)^\lambda dxight)$$

$$\leq \left(\int_{R^n} e^{-\langle Ax, x \rangle} dx\right)^{1-\lambda} \left(\int_{R^n} e^{-\langle Bx, x \rangle} dx\right)^\lambda$$

$$= \frac{(2\pi)^{n/2}}{(\det A)^{\frac{1-\lambda}{2}} (\det B)^{\frac{\lambda}{2}}}$$

and the lemma follows.

End of detour, back to main text.

The different ellipsoids connected with a convex body (or the different positions, corresponding to different choices of a Euclidean structure) that we shall consider in this chapter reflect different traces of symmetries which the convex body has. In some special cases, when the body has enough symmetries, meaning that there is a unique ellipsoid that satisfies all these symmetries, most (or all) of these positions coincide. This is the case, for example, when the normed space has a *symmetric basis* (i.e., a basis with respect to which the norm is invariant under permutations of the coordinates and changes of signs in the coordinates). However, in the general case a convex body has several different, useful, positions.

The most basic and classical positions of convex bodies which arise as solutions of extremal problems. These include John position (also called maximal volume ellipsoid position, minimal surface area position, minimal mean width position. The ideas behind the proofs are quite similar in all of these (so we might discuss only the John position) and is that when a position
is extremal then some differential must vanish, and its vanishing is connected with isotropicity (a term to be explained below) of some connected measure.

We shall then consider some direct applications of John position, and introduce a main tool, which is useful in many other results in the theory, the Brascamp-Lieb inequality.

Since the course is short and we will probably not touch upon other positions, let me mention briefly two other very important positions which are even more useful in the field.

The first, which is arose from classical mechanics back in the 19th century, is the isotropic position. It is defined to be the (unique, up to orthogonal transformations) position in which the body has a centre of mass at the origin, the inertia matrix of the convex body is scalar and the volume of the body is one. Namely, given a convex body $K$ in $\mathbb{R}^n$, the isotropic position of $K$ is defined as $\tilde{K} = T(K) + x_0$ with $T \in GL_n$ so that $\text{Vol}_n(\tilde{K}) = 1$, $\int_{\tilde{K}} x \, dx = 0$ and

$$\int_{\tilde{K}} \langle x, \theta \rangle^2 \, dx = L_K^2$$

for any $\theta \in S^{n-1}$, where $L_K > 0$ is a constant called the isotropic constant of $K$. The isotropic position turns out to have a characterization as a solution of an extremal problem: indeed, the isotropic position $\tilde{K} = T(K)$ of $K$ minimizes the quantity $\int_{\tilde{K}} |x|^2 \, dx$ over all $T \in GL_n$ such that $\text{Vol}_n(\tilde{K}) = 1$ and $\int_{\tilde{K}} x \, dx = 0$.

The second position which is very useful in many results and which dates to the mid 1980’s is the so called Milman position, or $M$-position of a convex body. It can be described “isometrically” as the position $\tilde{K} = x_0 + T(K)$ with the properties that $\text{Vol}_n(\tilde{K}) = \text{Vol}_n(B^*_2)$ and $\int_{\tilde{K}} x \, dx = 0$, which in addition minimizes the volume $\text{Vol}_n(\tilde{K} + B^*_2)$ among all such $T \in GL_n$. However, such a characterization hides its main properties and advantages that are in fact of an “isomorphic” nature. It turns out that the ellipsoid connected with the $M$-position (which is called the Milman ellipsoid or in short $M$-ellipsoid of $K$) can replace the original body in volume estimates. Given a body that is in $M$-position, the covering numbers of the body by a Euclidean ball and of a Euclidean ball by the body are not large. Furthermore, not only does this hold for the body, but also for its dual. Each of the two above facts imply that, for any pair of convex bodies $K$ and $T$ that are in $M$-position, a reverse Brunn-Minkowski type inequality is true, namely there exists a universal constant $C$ such that for any $n$, and any $K,T \in \mathbb{K}^n$ that are in $M$-position, one has

$$\text{Vol}_n(K + T)^{1/n} \leq C \left[ \text{Vol}_n(K)^{1/n} + \text{Vol}_n(T)^{1/n} \right].$$

One reason to introduce positions is to talk about distances between normed spaces up to isometry, so we consider equivalence classes of convex bodies “up to position”. To this end let us define

**Definition 6.3** (Banach-Mazur distance). Let $X = (\mathbb{R}^n, \| \cdot \|_K)$ and $Y = (\mathbb{R}^n, \| \cdot \|_T)$. The Banach-Mazur distance between $X$ and $Y$, denoted $d_{BM}(X,Y)$ is the infimal constant $C > 0$ such that there exists $A : X \to Y$ satisfying

$$\|x\|_K \leq \|Ax\|_T \leq C\|x\|_K$$

for all $x \in X$. Equivalently we have $A(K) \subset T \subset CA(K)$. For simplicity we will usually write
Given two centrally symmetric convex bodies $K$ and $T$ in $\mathbb{R}^n$ we agree to write

$$d(K, T) = d_{BM}(K, T) := d_{BM}(X_K, X_T) = \min\{t > 0 : A(K) \subseteq T \subseteq tA(K) \text{ for some } A \in GL(n)\}.$$ 

We also extend the definition of the Banach-Mazur distance to the setting of not necessarily symmetric convex bodies as follows: if $K$ and $T$ are convex bodies in $\mathbb{R}^n$ then the distance of $K$ and $T$ is defined by

$$d_g(K, T) = d_{BM}(K, T) = \min\{t > 0 : A(K) + y \subseteq T + x \subseteq t(A(K) + y)\}$$

where the minimum is over all $x, y \in \mathbb{R}^n$ and $A \in GL(n)$. Finally, it will also be useful to have the notion of the geometric distance between two spaces (or bodies), which is denoted $d_g(K, T)$, and defined as follows

$$d_g(K, T) = \inf\{ab : a, b > 0, \frac{1}{b}T \subset (K - x_0) \subset aT\}.$$ 

Of course we then have

$$d_{BM}(K, T) = \inf\{d_g(K, AT) : A \in GL(n)\}.$$

The main result in this section will show that for any $n$ dimensional normed space, $d_{BM}(X_K, \ell^n_2) \leq \sqrt{n}$. To do so, given a centrally symmetric convex body $K$ we shall take the ellipsoid $E$ of maximal volume inside $K$, and show that $K \subseteq \sqrt{n}E$. A similar result holds in the non-symmetric case with $n$ instead of $\sqrt{n}$ but we shall not go into it.

### 6.1 John's theorem

We shall show that there is a unique ellipsoid $E$ of maximal volume that is contained in $K$. We will say that $E$ is the maximal volume ellipsoid of $K$.

John’s theorem gives a characterization of the maximal volume ellipsoid, or, more precisely, a characterization for a body being in “John position”.

**Definition 6.4.** Given a compact convex body $K \subset \mathbb{R}^n$ that has the origin as an interior point we shall say that it is in “John position” if $B^n_2$ is a maximal volume ellipsoid of $K$.

It is easy to see by standard compactness arguments (which we illustrate below for completeness) that every such $K$ has some linear image $A(K)$ which is in John position. We shall prove below that the maximal volume ellipsoid is unique, and thus the John position of a convex body is unique up to orthogonal transformations. One of the most interesting applications of John position is the fact that, for any centrally symmetric convex body $K$, we can find an ellipsoid $E$ such that

$$E \subseteq K \subseteq \sqrt{n}E.$$ 

In particular, if $K$ is in John position, then $E = B^n_2$. In the language of geometric functional analysis this means that the Banach-Mazur distance between any $n$ dimensional normed space and $\ell^n_2$ is at most $\sqrt{n}$.

We may now reformulate (14) as $d(X, \ell^n_2) \leq \sqrt{n}$ for any $X = (\mathbb{R}^n, \|\cdot\|_K)$. Since this fact has an elementary proof we shall devote the next subsection to it. We shall then discuss minimal
and maximal volume ellipsoids of a convex body. Finally, we shall state and prove John’s theorem regarding the contact points of $K$ and $B^n_2$ when $K$ is in John position.

6.1.1 The “simple” John’s theorem

**Theorem 6.5** (John). Let $K$ be a centrally symmetric convex body in John position. Then $K \subseteq \sqrt{n}B^n_2$.

**Proof.** Assume towards a contradiction that there exists some $x_0 \in K$ with $|x_0| = R > \sqrt{n}$. Without loss of generality assume that $x_0 = Re_1$. From symmetry, $\text{conv}(\pm Re_1, B^n_2) \subseteq K$.

Consider an ellipsoid
\[
E = \left\{ x : \frac{x_1^2}{a^2} + \sum_{i=2}^{n} \frac{x_i^2}{b_i^2} \leq 1 \right\},
\]
where we set $b_2 = (1 - \varepsilon)$ for a small $\varepsilon$, and $a > 1$ is to be determined later in such a way that we will have $E \subseteq K$. The volume of $E$ is $ab^{n-1}\kappa_n$. To have it included in the aforementioned convex hull, note that we need just a two-dimensional argument, since we are working with revolution bodies. We shall need the following two drawings: the first is the picture of the convex hull, the second is the same picture, squeezed in one axis by $b/a$.

We see that $1/R = t/\sqrt{R^2 + t^2}$ from the first picture. Thus $t^2R^2 = R^2 + t^2$ and $t^2 = R^2/(R^2 - 1)$.

From the second picture we see that $a/R = b/(Rb/a) = t/\sqrt{R^2b^2/a^2 + t^2}$. That is, $R^2b^2 + t^2a^2 = R^2t^2$, and by replacing $t^2$ with $R^2/(R^2 - 1)$ we get
\[
a^2 = b^2 + R^2(1 - b^2).
\]

Writing also $1 - \varepsilon$ instead of $b^2$, we see that $a = 1 + (R^2 - 1)\varepsilon$ and that the square of the volume of the whole ellipsoid is equal to
\[
(1 - \varepsilon)^{n-1}(1 + \varepsilon(R^2 - 1)) = (1 - (n - 1)\varepsilon + o(\varepsilon))(1 + (R^2 - 1)\varepsilon) = 1 + (R^2 - n)\varepsilon + o(\varepsilon).
\]

Therefore, if it held that $R > \sqrt{n}$, then, for small enough $\varepsilon$, this volume would indeed be greater than 1, a contradiction.

6.1.2 Maximal and minimal volume ellipsoids

The fact that there cannot be two different maximal volume ellipsoids inside a convex body follows easily from the following result on determinants of matrices (this result goes back to
Minkowski).

**Proposition 6.6.** Given a convex body \( K \subset \mathbb{R}^n \), there exists a unique ellipsoid of maximal volume inscribed in \( K \).

**Proof.** The existence of such an ellipsoid follows easily from standard compactness arguments in \( \mathbb{R}^{n(n+1)/2} \), or from Blaschke’s selection theorem. Next assume that there are two different ellipsoids of maximal volume inside \( K \). Without loss of generality one of them is \( B_2^n \), and the other can be written as \( x_0 + A(B_2^n) \), with \( A \) a positive definite matrix which then must satisfy \( \det(A) = 1 \). By convexity of \( K \) we also have

\[
\frac{x_0}{2} + \frac{A + \text{Id}}{2} B_2^n \subseteq K.
\]

This is an ellipsoid, whose volume is equal to \( \det \left( \frac{A + \text{Id}}{2} \right) \kappa_n \), with the last determinant being at least 1 by Lemma 6.2. The equality condition implies that \( A = \text{Id} \) which means that \( K \) contains \( B_2^n \) and a translate of it, \( x_0 + B_2^n \) for some \( x_0 \neq 0 \) (as we assumed the ellipsoids were different). But then it contains their convex hull, which is easily seen to contain a slightly more elongated ellipsoid, and we get a contradiction to \( B_2^n \) having the maximum volume among all ellipsoids in \( K \).

Dually there is a unique ellipsoid \( E \) which contains a convex body \( K \) in and has minimal volume (the **minimal volume ellipsoid** of \( K \), you will show this in the exercise sheet.

### 6.1.3 Contact points and John’s theorem

Assume that a convex body \( K \) in \( \mathbb{R}^n \) that contains 0 in its interior is in John position. We will say that \( x \in \mathbb{R}^n \) is a contact point of \( K \) and \( B_2^n \) if \( |x| = \|x\|_K = \|x\|_{K^o} = 1 \). Note that under the assumption that \( B_2^n \subseteq K \) the first two conditions \( |x| = \|x\|_K = 1 \) imply the third one \( \|x\|_{K^o} = 1 \) since both \( B_2^n \) and \( K \) have \( x + x^\perp \) as a supporting hyperplane at \( x \) (indeed, the former conditions imply that any supporting hyperplane of \( K \) at \( x \) must also be a supporting hyperplane of \( B_2^n \) at \( x \), and thus this can only be \( x + x^\perp \)). John’s theorem describes the distribution of contact points on the unit sphere \( S^{n-1} \).

**Theorem 6.7** (John). If \( B_2^n \) is the maximal volume ellipsoid of a centrally symmetric convex body in \( \mathbb{R}^n \), there exist contact points \( x_1, \ldots, x_m \) of \( K \) and \( B_2^n \), and positive real numbers \( c_1, \ldots, c_m \) such that

\[
x = \sum_{j=1}^{m} c_j \langle x, x_j \rangle x_j
\]

for every \( x \in \mathbb{R}^n \). Moreover, one may choose \( m \leq \left( \frac{n+1}{2} \right) + 1 \).

**Remark 6.8.** Theorem 6.7 says that the identity operator \( \text{Id} \) of \( \mathbb{R}^n \) can be represented in the form

\[
\text{Id} = \sum_{j=1}^{m} c_j x_j \otimes x_j,
\]

where \( x_j \otimes x_j \) is the projection in the direction of \( x_j \): \( (x_j \otimes x_j)(x) = \langle x, x_j \rangle x_j \). Note that, if such a representation exists, then for every \( x \in \mathbb{R}^n \)

\[
|x|^2 = \langle x, x \rangle = \sum_{j=1}^{m} c_j \langle x, x_j \rangle^2.
\]
Also, if we choose \( x = e_i, i = 1, \ldots, n \), where \( \{e_i\} \) is the standard orthonormal basis of \( \mathbb{R}^n \), we have
\[
\sum_{i=1}^{n} |e_i|^2 = \sum_{i=1}^{n} \sum_{j=1}^{m} c_j \langle e_i, x_j \rangle^2 = \sum_{j=1}^{m} c_j \sum_{i=1}^{n} \langle e_i, x_j \rangle^2
\]
\[
= \sum_{j=1}^{m} c_j |x_j|^2 = \sum_{j=1}^{m} c_j.
\]

**Proof of Theorem 6.7.** The preceding remark shows that, if such a representation exists, we must have \( \sum_{j=1}^{m} \frac{c_j}{n} = 1 \). Our purpose is then to show that \( \text{Id}/n \) is a convex combination of matrices of the form \( x \otimes x \), where \( x \) is a contact point of \( K \) and \( B_2^n \). To this end, we define
\[
\mathcal{C} = \{ x \otimes x : |x| = \|x\|_K = 1 \}
\]
and show that \( \text{Id}/n \in \text{conv}(\mathcal{C}) \). Note that \( \text{conv}(\mathcal{C}) \) is a non-empty compact convex subset of \( \mathbb{R}^{n^2} \) (actually, of the cone of positive semi-definite matrices, which is a subset of an \( \binom{n+1}{2} \))-dimensional subspace of \( \mathbb{R}^{n^2} \).

Let us assume towards contradiction that \( \text{Id}/n \) may be separated from \( \text{conv}(\mathcal{C}) \). Then we assume that there is a linear functional \( \phi \) such that
\[
\langle \phi, \text{Id}/n \rangle < r \leq \langle \phi, u \otimes u \rangle
\]
for all \( u \in S^{n-1} \cap \partial K \) and some \( r \in \mathbb{R} \). We may also assume without loss of generality that \( \phi = \phi^* \) since otherwise we can set \( \psi = \frac{\phi + \phi^*}{2} \) and then, because both \( u \otimes u \) and \( \text{Id}/n \) are symmetric, \( \psi \) will satisfy the same inequality. (Equivalently, one may right away restrict to the space of linear symmetric matrices which is isometric to its dual space, and which \( \text{Id}/n \) as well as the elements of \( \mathcal{C} \) belong to, and separate there.)

Note that the traces of both \( u \otimes u \) and of \( \text{Id}/n \) are 1, so that the “scalar product” with the identity matrix is equal for both. In other words, we may subtract from \( \phi \) any linear functional that assigns to each matrix its trace times any constant \( c \) as a value, and still get separation, with \( r \) changed to \( r - c \). If this constant is chosen to be \( \text{tr}(\phi)/n \) we get that the new functional takes the value 0 at \( \text{Id}/n \).

So far we have shown that there exists a linear functional \( B \), given by a symmetric matrix \( B = B^* \), such that \( \langle B, \text{Id} \rangle = \text{tr}(B) = 0 \) and \( \langle B, u \otimes u \rangle > s > 0 \) for all \( u \in S^{n-1} \cap \partial K \).

The idea is to use \( B \) in order to define an ellipsoid of greater volume than \( B_2^n \) which still resides in \( K \); this ellipsoid shall be defined as
\[
\mathcal{E}_\delta = \{ x : \langle (\text{Id} + \delta B)x, x \rangle \leq 1 \}
\]
for a suitably chosen \( \delta \). Note that for small enough \( \delta \) the above set is indeed an ellipsoid as \( \text{Id} + \delta B \) is still positive definite (simply take \( \delta < 1/\|B\| \) (where \( \|B\| = \max\{|B(x)| : x \in S^{n-1}\} \)) and then \( \text{Id} + \delta B \) will have a square root, say \( S_\delta \), so that \( \mathcal{E}_\delta = S_\delta^{-1}(B_2^n) \).

Next, we shall show two things: first, that for small enough \( \delta \) this ellipsoid is indeed inside \( K \), and second, that the volume of this ellipsoid is greater than the volume of \( B_2^n \) (because the trace of \( B \) is 0). To show the former, we first consider the set of contact points of \( K \) and \( B_2^n \); denote this set by \( U \). We separate the points of \( S^{n-1} \) to points that are close to \( U \) (namely points \( v \in S^{n-1} \) with \( \text{dist}(v, U) \leq s/(2\|B\|) \)) and points that are far from it.
Let us look first at points of $S^{n-1}$ which are far from contact points: $V = \{v \in S^{n-1} : \text{dist}(v,U) \geq s/(2\|B\|)\}$. This is a compact set and thus it has positive distance from the boundary of $K$; in particular, $\alpha = \max\{\|x\| : x \in V\} < 1$. We claim that, for small enough $\delta$, $v/\|v\| \not\in \mathcal{E}_\delta$ for all $v \in V$ (geometrically this means that $\mathcal{E}_\delta \cap \mathbb{R}^+ v \subseteq K$ for all $v \in V$).

Indeed, let $\lambda = \min_{v \in V} \langle Bv, v \rangle$. Note that this is a negative number: since the trace of $B$ is zero (and given that the quantity $\langle Bx, x \rangle$ is not identically zero on $\mathbb{R}^n$), there must be some $w$ such that $\langle Bw, w \rangle < 0$. However, for all $u \in U$ we have $\langle Bu, u \rangle > s > 0$ and hence we must also have

$$\langle Bv, v \rangle = \langle Bu, u \rangle + \langle B(v-u), v \rangle > s - 2\|v-u\|\|B\| > 0$$

for every $v$ such that $|v-u| < s/(2\|B\|)$. Therefore, $w \in V$ and we conclude that $\lambda \leq \langle Bw, w \rangle < 0$.

Now is the time to choose $\delta < (1 - \alpha^2)/|\lambda|$. Then for $v \in V$

$$\left\langle (\text{Id} + \delta B)\left(\frac{v}{\|v\|}\right), \frac{v}{\|v\|}\right\rangle = \frac{1 + \delta \langle Bv, v \rangle}{\|v\|^2} \geq \frac{1 + \delta \lambda}{\alpha^2} > 1.$$

This completes the assertion in the case of $V$.

For $S^{n-1} \setminus V$ the same assertion follows easily by our choice of neighborhood of $U$: for every $u \in U$ we know that $\langle Bu, u \rangle = \langle B, u \otimes u \rangle > s$, and thus $\langle (\text{Id} + \delta B)u, u \rangle > 1 + \delta s$. Thus $u \not\in \mathcal{E}_\delta$. Furthermore, if $v \in S^{n-1}$ then

$$|\langle (\text{Id} + \delta B)v, v \rangle - \langle (\text{Id} + \delta B)u, u \rangle| = \delta|\langle Bv, v \rangle - \langle Bu, u \rangle| \leq \delta|\langle Bv, v \rangle - \langle Bu, u \rangle| + \delta|\langle Bu, u \rangle - \langle Bu, u \rangle| \leq 2\delta\|B\||u-v|.$$

This shows that for an $s/(2\|B\|)$-neighbourhood of $U$ on $S^{n-1}$, we have $\langle (\text{Id} + \delta B)v, v \rangle > 1$ and in particular $v \not\in \mathcal{E}_\delta$. However $v \in B_2^n \subseteq K$, and hence $\mathcal{E}_\delta \cap \mathbb{R}^+ v \subseteq K$ for all $v \in S^{n-1} \setminus V$ too.

We have completed the first part, namely to show that $\mathcal{E}_\delta \subseteq K$ for small enough $\delta$. As for the volume of this ellipsoid, we have

$$\text{Vol}(\mathcal{E}_\delta) = \kappa_n / \det(\text{Id} + \delta B)^{1/2}.$$

Note now that, since $\det(\text{Id} + \delta B)^{1/n} \leq \text{tr}(\text{Id} + \delta B)/n = 1$ by the arithmetic-geometric means inequality and since $B_2^n$ has the maximum volume among the ellipsoids contained in $K$, we ought to have that the ellipsoid $\mathcal{E}_\delta$ has the same volume as $B_2^n$, and hence that it coincides with the maximal volume ellipsoid in $K$ which, by our assumption, is $B_2^n$. We thus arrive at a contradiction since $B \neq 0$. We conclude that $\text{Id}/n \in \text{conv}(\mathcal{C})$. In addition, by Carathéodory’s...
theorem from classical convexity, applied to the set $C$ which is a subset of an $(n+1)/2$-dimensional subspace of $\mathbb{R}^{n^2}$, we may write $\text{Id}/n$ as a combination of at most $m = (n+1)/2 + 1$ points in $\text{conv}(C)$.

**Definition 6.9.** A Borel measure $\mu$ on $S^{n-1}$ is called isotropic if

$$\int_{S^{n-1}} \langle x, \theta \rangle^2 d\mu(x) = \frac{\mu(S^{n-1})}{n}$$

for every $\theta \in S^{n-1}$. We will make frequent use of the next standard lemma.

**Lemma 6.10.** Let $\mu$ be a Borel measure on $S^{n-1}$. The following are equivalent:

(i) $\mu$ is isotropic.

(ii) For every $i, j = 1, \ldots, n$,

$$\int_{S^{n-1}} \phi_i \phi_j d\mu(\phi) = \frac{\mu(S^{n-1})}{n} \delta_{i,j}.$$

(iii) For every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$,

$$\int_{S^{n-1}} \langle \phi, T\phi \rangle d\mu(\phi) = \frac{\text{tr}(T)}{n} \mu(S^{n-1}).$$

**Proof.** Setting $\theta = e_i$ and $\theta = \frac{e_i + e_j}{\sqrt{2}}$ in (15) we get (16). Moreover, from the observation that if $T = (t_{ij})_{i,j=1}^n$ then $\langle \phi, T\phi \rangle = \sum_{i,j=1}^n t_{ij} \phi_i \phi_j$, we readily see that (16) implies (17). Finally, note that applying (17) with $T(\phi) = \langle \phi, \theta \rangle \theta$ we get (15).

**Note.** Theorem 6.7 implies that

$$\sum_{j=1}^m c_j \langle x_j, \theta \rangle^2 = 1$$

for every $\theta \in S^{n-1}$. In our terminology, the measure $\mu$ on $S^{n-1}$ that gives mass $c_j$ to the point $x_j$, $j = 1, \ldots, m$, is isotropic.

Conversely, we have the following proposition which shows that if a body $K$ that contains $B^n_2$ has enough points in $\partial K \cap S^{n-1}$, enough meaning that an isotropic measure with support on this set can be constructed, then the body must be in John position. This is very useful in that we can sometimes immediately determine whether a body is in John position by observing the distribution of its contact points.

**Proposition 6.11 (Ball).** Let $K$ be a convex body in $\mathbb{R}^n$ that contains the Euclidean unit ball $B^n_2$. Assume that the set of contact points of $K$ and $B^n_2$ is non-empty and assume that there exists an isotropic Borel measure $\mu$ on $S^{n-1}$ whose support lies in this set. Then, $B^n_2$ is the maximal volume ellipsoid of $K$.

**Proof.** Instead of $K$ consider the set

$$L = \{ y : \langle x, y \rangle \leq 1 \text{ for all } x \in \text{supp}(\mu) \}.$$  

(Note: this is the polar of conv(supp($\mu$))). Clearly $K \subseteq L$ (since $\|x\|_{K^*} = 1$ for every contact point $x$ of $K$) and so it is enough to show that $B^n_2$ is maximal for $L$. Pick some ellipsoid $E \subseteq L$
given by
\[ \mathcal{E} = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^{n} \frac{\langle x, v_i \rangle^2}{\alpha_i^2} \leq 1 \right\}, \]
where \( \{v_i : i = 1, \ldots, n\} \) is some orthonormal basis and \( \alpha_i \) are the respective lengths of the semi-axes.

If \( u \in \text{supp}(\mu) \) then the vector \( y \equiv y(u) \) defined as
\[ y(u) = \sum_{j=1}^{n} \alpha_j \langle u, v_j \rangle v_j \]
belongs to \( \mathcal{E} \subseteq L \). Indeed, \( y(u) \) has “\( j \)-th coordinate” (that is, with respect to the basis \( \{v_i\} \)) equal to \( \alpha_j \langle u, v_j \rangle \), so that the sum of the squares weighted by \( \alpha_j^{-2} \) is \( |u|^2 = 1 \). Since \( y(u) \in \mathcal{E} \subseteq L \), we see that \( \langle u, y(u) \rangle \leq 1 \), which gives \( \sum_{j=1}^{n} \alpha_j \langle u, v_j \rangle^2 \leq 1 \). This holds for all \( u \in \text{supp}(\mu) \).

Isotropicity means that for every \( x \),
\[ \int_{S^{n-1}} \langle x, u \rangle^2 d\mu(u) = |x|^2 \mu(S^{n-1})/n. \]
Using this for the \( v_j \), we may now write
\[ \mu(S^{n-1}) = \int_{S^{n-1}} d\mu(u) \geq \int_{S^{n-1}} \langle y(u), u \rangle d\mu(u) \]
\[ = \int_{S^{n-1}} \sum_{j=1}^{n} \alpha_j \langle v_j, u \rangle^2 d\mu(u) = \frac{\mu(S^{n-1})}{n} \sum_{j=1}^{n} \alpha_j. \]

By the arithmetic-geometric means inequality, \( \sum_{j=1}^{n} \alpha_i \leq n \) implies \( \prod_{j=1}^{n} \alpha_i \leq 1 \), or, in other words, that \( \text{Vol}_n(\mathcal{E}) \leq \kappa_n \) as claimed.

**Remark 6.12.** John’s theorem about the distribution of contact points still holds for a general convex body \( K \subseteq \mathbb{R}^n \); namely we have:

**Theorem.** Let \( K \) be a convex body in \( \mathbb{R}^n \). The Euclidean unit ball \( B_2^n \) is the ellipsoid of maximal volume in \( K \) if and only if \( K \) contains \( B_2^n \) and there is a sequence \( (x_j)_{j=1}^{m} \) of contact points of \( B_2^n \) and \( \partial K \) and a sequence \( (c_j)_{j=1}^{m} \) of positive numbers so that
\[ \sum_{j=1}^{m} c_j x_j = 0 \]
and
\[ \text{Id} = \sum_{j=1}^{m} c_j x_j \otimes x_j. \]

Verify that the proof carries over to this case as well.

Finally let us show how Theorem 6.7 implies Theorem 6.5 above which states that, if \( K \) is a centrally symmetric convex body in \( \mathbb{R}^n \) and \( \mathcal{E} \) is the maximal volume ellipsoid of \( K \), then \( K \subseteq \sqrt{n} \mathcal{E} \). This is equivalent to the statement that, if \( B_2^n \) is the maximal volume ellipsoid of \( K \), then \( K \subseteq \sqrt{n} B_2^n \).
Proof of Theorem 6.5. Consider the representation of the identity

\[ x = \sum_{j=1}^{m} c_j \langle x, x_j \rangle x_j \]

of Theorem 6.7. We will use the fact that \( \|x_j\|_K = \|x_j\|_{K^*} = |x_j| = 1, j = 1, \ldots, m \). For every \( x \in K \) we have

\[ |x|^2 = \sum_{j=1}^{m} c_j^2 \langle x, x_j \rangle^2 \leq \sum_{j=1}^{m} c_j (\|x\|_K \|x_j\|_{K^*})^2 \leq \sum_{j=1}^{m} c_j = n. \]

This shows that \( |x| \leq \sqrt{n} \). Therefore, \( B^n_2 \subseteq K \subseteq \sqrt{n}B^n_2 \).

6.1.4 Contact points and the Dvoretzky-Rogers lemmas

Let \( X = (\mathbb{R}^n, \| \cdot \|) \) be an \( n \)-dimensional normed space. Assume that \( B^n_2 \) is the ellipsoid of maximal volume contained in the unit ball \( K \) of \( X \). Recall that from John’s theorem there exist \( x_1, \ldots, x_m \in \mathbb{R}^n \) with \( \|x_j\| = |x_j| = \|x_j\|_* = 1 \) and \( c_1, \ldots, c_m > 0 \) such that \( \text{Id} = \sum_{j=1}^{m} c_j x_j \otimes x_j \). In other words, for every \( x \in \mathbb{R}^n \),

(18)

\[ x = \sum_{j=1}^{m} c_j \langle x, x_j \rangle x_j. \]

From (18) it follows that

\[ \sum_{j=1}^{m} c_j \langle x, x_j \rangle^2 = |x|^2, \]

and thus, by applying this with \( x = e_i \) for every \( i = 1, \ldots, n \) and summing, that \( \sum_{j=1}^{m} c_j = n. \)

One may assume that \( m \leq \left( \frac{n+1}{2} \right) + 1. \)

In this subsection we collect several useful results on the distribution of the contact points \( x_j \) on \( S^{n-1} \), known as “Dvoretzky-Rogers type lemmas”.

Lemma 6.13. If \( B^n_2 \) is the maximal volume ellipsoid of a convex body \( K \in \mathcal{K}_n^{(0)} \), then for every \( T \in L(\mathbb{R}^n) \) we can find a contact point \( y \) of \( K \) and \( B^n_2 \) with the property

\[ \langle y, Ty \rangle \geq \frac{\text{tr}(T)}{n}. \]

Proof. We have

\[ \text{tr}T = \langle T, \text{Id} \rangle = \sum_{j=1}^{m} c_j \langle T, x_j \otimes x_j \rangle. \]

Since \( \sum_{j=1}^{m} c_j = n \), we may find \( y \) among the \( x_j \) that satisfies

\[ \langle y, Ty \rangle = \langle T, y \otimes y \rangle \geq \frac{\text{tr}(T)}{n} \]

as desired. \( \square \)
Theorem 6.14 (Dvoretzky-Rogers).\ If $B^n_2$ is the maximal volume ellipsoid of a convex body $K \in \mathcal{K}^n(0)$, there exists an orthonormal sequence $z_1, \ldots, z_n$ in $\mathbb{R}^n$ such that

$$\left(\frac{n-i+1}{n}\right)^{1/2} \leq \|z_i\| \leq |z_i| = 1$$

for all $i = 1, \ldots, n$.

Proof. We define the $z_i$ inductively. Choose $z_1$ to be any contact point of $K$ and $B^n_2$, and assume that $z_1, \ldots, z_k$ have been chosen for some $k < n$.

We set $F_k = \text{span}\{z_1, \ldots, z_k\}$. Then, $\text{tr}(P_{F_k^\perp}) = n - k$, and applying Lemma 6.13 we may find a contact point $y_{k+1}$ of $K$ and $B^n_2$ with

$$|P_{F_k^\perp}y_{k+1}|^2 = \langle y_{k+1}, P_{F_k^\perp}y_{k+1} \rangle \geq \frac{n-k}{n}.$$ 

It follows that $\|P_{F_k^\perp}y_{k+1}\| \leq |P_{F_k^\perp}y_{k+1}| \leq \sqrt{k/n}$.

We define $z_{k+1} = P_{F_k^\perp}y_{k+1}/|P_{F_k^\perp}y_{k+1}|$. Then, since $h_K(y_{k+1}) = 1$, we have

$$1 = |z_{k+1}| \geq \|z_{k+1}\| \geq \langle y_{k+1}, z_{k+1} \rangle = \left|P_{F_k^\perp}y_{k+1}\right| \geq \left(\frac{n-k}{n}\right)^{1/2},$$

and the inductive step is complete. \qed

Corollary 6.15. Assume that $B^n_2$ is the maximal volume ellipsoid of a convex body $K \in \mathcal{K}^n(0)$. If $k = \lfloor n/2 \rfloor + 1$, we can find orthonormal vectors $z_1, \ldots, z_k$ such that

$$\frac{1}{\sqrt{2}} \leq \|z_j\| \leq 1$$

for all $j = 1, \ldots, k$.

Remark 6.16. It is useful to note that one also has a version of Corollary 6.15 that provides a bound for the norms of all of the orthonormal vectors $z_i$, I left it for you to prove in the exercise sheet.

6.2 Another proof for John’s theorem

As explained in class, in a maximization problem we can many times deduce the existence of some good structure (special contact points, an isotropic measure) simply by using the fact that a certain differential vanishes. Let us have a second look at John’s theorem with these “eyeglasses” on.

I am doing the centrally symmetric case for simplicity.

Lemma 6.17. Let $K = -K$ be a convex body and assume $B^n_2 \subset K$ is the maximal volume ellipsoid. Then for any $T \in SL_n$ there exists a point $u \in S^{n-1}$ such that

$$\|Tu\|_K \geq 1.$$ 

Proof. Since the maximal volume position is unique, $TB^n_2$ is not a subset of $K$ (unless $T$ is orthonormal, in which case any contact point $v$ would provide $u = T^{-1}v \in S^{n-1}$), and in this case there is some $Tu \not\in K$. \qed
Lemma 6.18. Let $K = -K$ be a convex body and assume $B^n_2 \subset K$ is the maximal volume ellipsoid. Then for any $S \in L(\mathbb{R}^n, \mathbb{R}^n)$ there exists a contact point $u \in S^{n-1} \cap \partial K$ such that

$$\|Su\|_K \geq \frac{1}{n} \text{tr} S.$$ 

Proof. Define $T_\varepsilon = \frac{1}{\det(I + \varepsilon S)^{1/n}} (I + \varepsilon S)$ so that $T_\varepsilon \in SL_n$. By Lemma 6.17 we may find for every $\varepsilon$ a point $u_\varepsilon \in S^{n-1}$ such that

$$\|T_\varepsilon u_\varepsilon\|_K \geq 1$$

that is, $\|u_\varepsilon + \varepsilon Su_\varepsilon\|_K \geq \det(I + \varepsilon S)^{1/n}$. Using that $B^n_2 \subset K$ we know $\|u_\varepsilon\| = |u_\varepsilon| = 1$ so that for $\varepsilon \geq 0$ we have

$$1 + \varepsilon \|Su_\varepsilon\|_K \geq \|Su_\varepsilon\|_K \geq \det(I + \varepsilon S)^{1/n} = 1 + \varepsilon \frac{1}{n} \text{tr} S + o(\varepsilon).$$

Therefore

$$\|Su_\varepsilon\|_K \geq \frac{1}{n} \text{tr} S + o(1).$$

Take now a sequence $u_\varepsilon \rightarrow u$ as $\varepsilon \rightarrow 0$ then $u \in S^{n-1}$ and $\|Su\|_K \geq \frac{1}{n} \text{tr} S$. Clearly as $1 \leq \|T_\varepsilon u_\varepsilon\|_K = \|u_\varepsilon + O(\varepsilon)\|_K$ we have that in the limit also $\|u\|_K \geq 1$ but since $B^n_2 \subset K$ we actually have $\|u\|_K = 1$, so that $u$ is a contact point.

Finally, we can show

Lemma 6.19. Let $K = -K$ be a convex body and assume $B^n_2 \subset K$ is the maximal volume ellipsoid. Then for any $T \in L(\mathbb{R}^n, \mathbb{R}^n)$ there exists a contact point $u \in S^{n-1} \cap \partial K$ such that

$$\langle Tu, u \rangle \geq \frac{1}{n} \text{tr} T.$$ 

Proof. Given $T$ define $S_\varepsilon = (I + \varepsilon T)$ and by Lemma 6.18 find for every $\varepsilon$ (now it may be also negative, by the way) a contact point $u_\varepsilon \in S^{n-1} \cap \partial K$ such that

$$\|u_\varepsilon + \varepsilon Tu_\varepsilon\|_K = \|S_\varepsilon u_\varepsilon\|_K \geq \frac{1}{n} \text{tr} S_\varepsilon = 1 + \varepsilon \frac{1}{n} \text{tr} T.$$ 

The left hand side is equal to

$$\|u_\varepsilon\| + \varepsilon \langle \nabla \|u_\varepsilon\|, Tu_\varepsilon \rangle + o(\varepsilon) = 1 + \varepsilon \langle \nabla \|u_\varepsilon\|, Tu_\varepsilon \rangle + o(\varepsilon).$$

The gradient of the norm is a functional $w \in \partial K^0$ which is the supporting functional at $u$, that is, it satisfies $\langle u, w \rangle = 1$. However, since $u$ is a contact point of $K$ and $B^n_2$, the normal at $u$ is $u$ itself. We thus get for every $\varepsilon$ a contact point $u_\varepsilon$ for which

$$1 + \varepsilon \langle u_\varepsilon, Tu_\varepsilon \rangle + o(\varepsilon) \geq 1 + \varepsilon \frac{1}{n} \text{tr} T,$$

which can be rewritten (for $\varepsilon > 0$, say) as

$$\langle u_\varepsilon, Tu_\varepsilon \rangle + O(\varepsilon) \geq \frac{1}{n} \text{tr} T.$$ 

Taking a converging subsequence of contact points as $\varepsilon \rightarrow 0^+$, we end up with a contact point.
Note that what we have just proven is Lemma 6.13 which in itself was useful for us to get Dvoretzky-Rogers.

Proof of John’s theorem. To complete the proof of John’s theorem in this vein is actually quite simple. Consider as in the original proof, the convex hull of contact points

\[ C = \{ u \otimes u : |u| = \|u\|_K = 1 \} \]

and assume it may be separated from \( \frac{1}{n} I \), using some linear functional given by \( T \in L(\mathbb{R}^n, \mathbb{R}^n) \).

This means that

\[ \langle T, \frac{1}{n} I \rangle > c > \langle T, u \otimes u \rangle \]

for any \( u \) which is a contact point. This means exactly

\[ \frac{\text{tr}T}{n} > c > \langle Tu, u \rangle \]

which contradicts Lemma 6.19.

7 Reverse isoperimetric inequality

Modulo affine transformations, among all convex bodies of a given volume in \( \mathbb{R}^n \), the \( n \)-dimensional simplex has “largest” surface area, while among centrally symmetric convex bodies, the cube is the extremal body.

**Theorem 7.1 (Ball).** Let \( K \) be a convex body in \( \mathbb{R}^n \) and \( T \) a regular \( n \)-dimensional solid simplex. Then there is an affine image \( \tilde{K} \) of \( K \) satisfying

\[ \text{Vol}(\tilde{K}) = \text{Vol}(T) \quad \text{and} \quad \text{Vol}_{n-1}(\partial(\tilde{K})) \leq \text{Vol}_{n-1}(\partial(T)). \]

If \( K \) is a centrally symmetric convex body in \( \mathbb{R}^n \) and \( Q \) an \( n \)-dimensional cube then there is an affine image \( \tilde{K} \) of \( K \) satisfying

\[ \text{Vol}(\tilde{K}) = \text{Vol}(Q) \quad \text{and} \quad \text{Vol}_{n-1}(\partial(\tilde{K})) \leq \text{Vol}_{n-1}(\partial(Q)). \]

A main tool in the proof of the above theorem is the Brascamp-Lieb inequality. In Section ?? we state and prove its “normalized form” put forward by K. Ball together with its reverse form, which is due to Barthe.

**Theorem 7.2 (Ball).** Let \( u_1, \ldots, u_m \in S^{n-1} \) and \( c_1, \ldots, c_m > 0 \) satisfy

\[ \text{Id} = \sum_{j=1}^{m} c_j u_j \otimes u_j. \]

If \( f_1, \ldots, f_m : \mathbb{R} \rightarrow \mathbb{R}^+ \) are measurable functions, then

\[ \int_{\mathbb{R}^n} \prod_{j=1}^{m} f_j^{c_j}(\langle u_j, x \rangle) dx \leq \prod_{j=1}^{m} \left( \int_{\mathbb{R}} f_j(t) dt \right)^{c_j} . \]
7.1 Maximal volume ratio

Definition 7.3. The volume ratio of a convex body $K$ is defined to be

$$\operatorname{vr}(K) = \inf_{E \subseteq K} \left( \frac{\operatorname{Vol}_n(K)}{\operatorname{Vol}_n(E)} \right)^{1/n}$$

where the infimum is taken over all ellipsoids $E$ inside $K$.

Theorem 7.4 (Ball). (i) Among centrally symmetric convex bodies in $\mathbb{R}^n$, the cube has the largest volume ratio.
(ii) Among convex bodies in $\mathbb{R}^n$, the simplex has the largest volume ratio.

Proof. (i) One has to show that if a centrally symmetric convex body $K$ is in John position then $\operatorname{Vol}_n(K) \leq 2^n$. But when a body is in John position, by Theorem 6.7 there exist contact points $u_j$ of it and $B^*_2$ and $c_j > 0$ such that

$$\operatorname{Id} = \sum_{j=1}^m c_j u_j \otimes u_j.$$

Since $K$ is symmetric, for every contact point $u_j$ of $K$ and $B^*_2$ we have that $-u_j$ is also a contact point. Thus the body is contained in the intersection of the strips $C := \{x : |\langle x, u_j \rangle| \leq 1\}$ (because for all contact points $u_j$ we have that $h_K(u_j) = 1$), and its volume is at most

$$\operatorname{Vol}_n(C) = \int_{\mathbb{R}^n} \prod_{j=1}^m 1_{[-1,1]}(\langle x, u_j \rangle)^{c_j} dx \leq \prod_{j=1}^m \left( \int_{\mathbb{R}} 1_{[-1,1]} \right)^{c_j} = 2^n$$

by the Brascamp-Lieb inequality and the fact that $\sum_{j=1}^m c_j = n$.

(ii) Moving to the non-symmetric case, we note that the $n$-dimensional simplex that circumscribes $B^*_2$ has volume

$$\frac{n^{n/2}(n + 1)(n+1)/2}{n!}$$

(it is easier to check this working in $\mathbb{R}^{n+1}$ on the hyperplane $\sum x_i = 1$, say). We thus need to establish this bound for the volume of a body in John position. Use John’s theorem in the non-symmetric case to find contact points $u_j$ and $c_j > 0$ such that $\sum_j c_j u_j = 0$ and $\operatorname{Id} = \sum_j c_j u_j \otimes u_j$. As above, the body $K$ lies inside the possibly larger body

$$L := \{x : \langle x, u_j \rangle \leq 1, i = j, \ldots, m\}$$

(this is a bounded set because $\sum_j c_j u_j = 0$), so it would suffice to bound this body’s volume by the desired bound.

We construct a new sequence of vectors $(v_i)_{i=1}^m$ in $\mathbb{R}^{n+1}$ which would be orthogonal in the extreme case where $K$ is a simplex. The estimate follows from an application of the Brascamp-Lieb inequality to a family of functions whose product is supported on a cone in $\mathbb{R}^{n+1}$ that has cross-sections similar to $K$: regard $\mathbb{R}^{n+1}$ as $\mathbb{R}^n \times \mathbb{R}$ and for each $j$, let

$$v_j = \sqrt{\frac{n}{n+1}} \left(-u_j, \frac{1}{\sqrt{n}} \right) \in \mathbb{R}^{n+1} \quad \text{and} \quad d_j = \frac{n+1}{n} c_j \in \mathbb{R}^+. $$
Direct computation shows that
\[ \text{Id}_{n+1} = \sum_{j=1}^{m} d_j v_j \otimes v_j. \]

Take \( f_j = e^{-t}, t \geq 0 \). Apply the Brascamp-Lieb inequality to get

\[
\int_{\mathbb{R}^{n+1}} \prod f_j^{d_j}(\langle x, v_j \rangle) dx \leq \prod_{j=1}^{m} \left( \int_{\mathbb{R}} f_j \right)^{d_j} = 1, \tag{19}
\]

because \( \int f_j = 1 \). Next, we compute the same integral over each hyperplane \( x_{n+1} = r \). One can check that the function \( \prod f_j^{d_j}(\langle x, v_j \rangle) \) is non-zero precisely when \( r \geq 0 \) and the point \( x \) is in \( \frac{r}{\sqrt{n}} L \times \{ r \} \) (given that each \( f_j \) is defined to be non-zero precisely at the non-negative \( t \in \mathbb{R} \)), and in that case it equals exactly \( e^{-r\sqrt{n+1}} \) (because we know that \( \sum_j c_j u_j = 0 \)). In other words,

\[
\int_{\{ x_{n+1} = r \}} \prod f_j^{d_j}(\langle x, v_j \rangle) dx = e^{-r\sqrt{n+1}} \text{Vol}_n \left( \frac{r}{\sqrt{n}} L \right) = e^{-r\sqrt{n+1}} \left( \frac{r}{\sqrt{n}} \right)^n \text{Vol}_n(L).
\]

Then, (19) implies

\[
1 \geq \text{Vol}_n(L) \int_{0}^{\infty} e^{-r\sqrt{n+1}} \left( \frac{r}{\sqrt{n}} \right)^n dr \geq \frac{\text{Vol}_n(L) \cdot n!}{n^{n/2(n+1)(n+1)/2}}.
\]

We conclude the proof noticing that \( \text{Vol}_n(K) \leq \text{Vol}_n(L) \) and that the upper bound for \( \text{Vol}_n(K) \) is no other than the number we had equality for in the case of the simplex. \( \square \)

### 7.2 Proof of the reverse isoperimetric inequality

It is perhaps surprising that in proving a statement about the maximal “minimal surface invariant”, we shall not use the minimal surface area position but rather the John position. Of course, in the case of the extremal bodies for the required inequality, these two positions coincide. We shall prove only the second part of Theorem 7.1, as the proof of the first part is identical.

**Proof of Theorem 7.1.** We have to find an affine image of a given body \( K \) such that

\[
\text{Vol}_{n-1}(\partial(\tilde{K})) \leq c_n \text{Vol}(\tilde{K})^{\frac{n-1}{n}}
\]

where \( c_n \) is determined so that for the (say, side-length-1) cube there is equality, that is, \( c_n = 2n \). This position will be the John position of \( K \): indeed, for this position we have

\[
\text{Vol}_{n-1}(\partial(K)) = \lim_{t \to 0} \frac{\text{Vol}_n(K + tB_2^n) - \text{Vol}_n(K)}{t} \\
\leq \lim_{t \to 0} \frac{\text{Vol}_n((1 + t)K) - \text{Vol}_n(K)}{t} = n \text{Vol}_n(K) \leq 2n \text{Vol}(K)^{\frac{n-1}{n}},
\]

66
where we make use of Theorem 7.4 for the last inequality.

For the simplex, we again take the body $K$ in John-position so that now

$$\Vol_{n-1}(\partial(K)) = \lim_{t \to 0} \frac{\Vol_n(K + tB^n_2) - \Vol_n(K)}{t} \leq \lim_{t \to 0} \frac{\Vol_n((1 + t)K) - \Vol_n(K)}{t} = n\Vol_n(K) = n\Vol(K)^{n-1/n} \Vol_n(K)^{\frac{1}{n}} \leq \frac{n^{3/2}(n + 1)^{(n+1)/2n}}{(n!)^{1/n}} \Vol(K)^{n-1},$$

where we make use of Theorem 7.4 for the last inequality. We should check that the simplex gives equality in this inequality, to this end we need to find the surface area of a simplex in John position, which is $(n+1)$ times the area of a simplex with side length which we computed in class; check. \hfill $\square$

### 7.3 Proof for Ball’s normalized form for the Brascamp-Lieb inequality and its inverse

We are now going to prove the Brascamp-Lieb theorem we used for the proof above, namely

**Theorem [Ball]** Let $u_1, \ldots, u_m \in S^{n-1}$ and $c_1, \ldots, c_m > 0$ satisfy

$$\text{Id} = \sum_{j=1}^m c_j u_j \otimes u_j.$$

If $f_1, \ldots, f_m : \mathbb{R} \to \mathbb{R}^+$ are measurable functions, then

$$\int_{\mathbb{R}^n} \prod_{j=1}^m f_j^{c_j}(\langle u_j, x \rangle) dx \leq \prod_{j=1}^m \left( \int_{\mathbb{R}} f_j(t) dt \right)^{c_j}.$$

As stated in class, the inequality is actually more general, and states that when you have some vectors $u_j$ and constants $p_j$ such that $\sum \frac{1}{p_j} = n$ then

$$\sup \frac{\int_{\mathbb{R}^n} \prod_{j=1}^m f_j^{c_j}(\langle u_j, x \rangle) dx}{\prod_{j=1}^m \left( \int_{\mathbb{R}} f_j^{p_j}(t) dt \right)^{1/p_j}}$$

is maximized for Gaussian functions $f_j$. The advantage of the normalization condition $\text{Id} = \sum_{j=1}^m c_j u_j \otimes u_j$ (where $c_j = 1/p_j$) is that the constant in the inequality, which is (always) attained by Gaussian functions, is 1.

Our method of proof will actually prove the theorem along with a reverse form, which we state in Section 7.3.3, and then prove both of them simultaneously. But before that, let examine some nice case and connect it to other things in life.

#### 7.3.1 Young inequality

Recall the usual convolution of two bounded integrable functions $f, g : \mathbb{R} \to \mathbb{R}$ is

$$f * g(x) = \int f(y)g(x-y)dy.$$
Young’s inequality states that if \( f \in L_p \) and \( g \in L_q \) and \( p^{-1} + q^{-1} = 1 + s^{-1} \) then
\[
\|f * g\|_s \leq \|f\|_p \|g\|_q.
\]

It actually holds on all locally compact groups (so \( \mathbb{R} \) is just one case), and the proof is just a bunch of cleverly applied Hölder inequalities. When the group is actually compact, like the circle, constant functions give an example of equality; But in \( \mathbb{R} \), there is no equality, and a sharper constant may be inserted. It was Beckner, and also Brascamp and Lieb, who showed that in \( \mathbb{R} \) the best constant is attained by Gaussians. [The actual precise constant is \( A_p A_q A_r \) where \( A_p = p^{1/p} / (p')^{1/p'} \).]

How is this connected to our previous discussions? Easy. to find \( \|f * g\|_s \) we need to take
\[
\sup \{ \int (f * g)h : h \in L_r, \|h\|_r = 1 \}
\]

where \( r^{-1} + s^{-1} = 1 \) (so that \( L_r \) is the dual space to \( L_s \)). Restating Young inequality is simply saying that if \( p^{-1} + q^{-1} + r^{-1} = 2 \) then
\[
\int (f * g)h = \int \int f(y)g(x - y)h(x)dydx \leq \|f\|_p \|g\|_q \|h\|_r.
\]

Define \( u_1 = e_2, u_2 = e_1 - e_2, u_3 = e_1 \) and this is very much what we wanted to talk about before. Of course, this is not a representation of identity exactly, but it is still in the form of the more general theorem stated above, and indeed the sharp constant in Young’s inequality is due to them.

We only use the normalized form in this text, so we stick to it from here onward.

### 7.3.2 Gaussian functions

We start with the technical fact, which is linear algebra. We quote here the Cauchy-Binet formula from linear algebra:

Let \( A \in M_{n \times m} \) and \( B \in M_{m \times n} \). For \( 1 \leq j_1, j_2, \ldots, j_n \leq m \) let \( A^{j_1 j_2 \ldots j_n} \) denote the \( n \times n \) matrix consisting of columns \( j_1, j_2, \ldots, j_n \) of \( A \) and \( B_{j_1 j_2 \ldots j_n} \) denote the \( n \times n \) matrix consisting of rows \( j_1, j_2, \ldots, j_n \) of \( B \). Then

\[
\det (AB) = \sum_{1 \leq j_1 < j_2 < \ldots < j_n \leq m} \det (A^{j_1 j_2 \ldots j_n}) \det (B_{j_1 j_2 \ldots j_n}).
\]

We shall use the formula as follows: Let \( \lambda_j > 0, j = 1, \ldots, m \). For every \( I \subseteq \{1, \ldots, m\} \) with cardinality \( |I| = n \) we define

\[
\lambda_I = \prod_{i \in I} \lambda_i \quad \text{and} \quad U_I = \left( \det \left( \sum_{j \in I} c_j u_j \otimes u_j \right) \right)^2.
\]

Then

**Lemma 7.5.** Under the above assumptions and notations, we have

\[
(20) \quad \det \left( \sum_{j=1}^m c_j \lambda_j u_j \otimes u_j \right) = \sum_{|I|=n} \lambda_I U_I.
\]
Proof. Let $A \in M_{n \times m}$ have columns $\lambda_j c_j u_j$ and let $B$ have rows $u_j$. Then

$$ABx = \sum_{j=1}^{m} \lambda_j c_j u_j \langle u_j, x \rangle.$$  

In other words, $AB = \left( \sum_{j=1}^{m} \lambda_j (\sqrt{c_j} u_j) \otimes (\sqrt{c_j} u_j) \right)$. By the Cauchy-Binet formula we have

$$\det \left( \sum_{j=1}^{m} c_j \lambda_j u_j \otimes u_j \right) = \sum_{|I|=n} \det \left( \sum_{j \in I} c_j \lambda_j u_j \otimes u_j \right).$$

However, we now simply have a diagonal matrix multiplying a square one, so that

$$\det \left( \sum_{j=1}^{m} c_j \lambda_j u_j \otimes u_j \right) = \sum_{|I|=n} \lambda_I \det U_I.$$

Next we give show that when restricted to one-dimensional Gaussian functions, the Brascamp-Lieb inequality holds with constant 1, and it is sharp (in other words, if we supermize the ratio in the Brascamp-Lieb theorem, over Gaussians, we get 1).

**Proposition 7.6.** Let $u_1, \ldots, u_m \in S^{n-1}$ and $c_1, \ldots, c_m > 0$ satisfy $\text{Id} = \sum_{j=1}^{m} c_j u_j \otimes u_j$. Then,

$$\sup \left\{ \frac{\int_{\mathbb{R}^n} \prod_{j=1}^{m} g_j^c (\langle x, u_j \rangle) dx}{\prod_{j=1}^{m} \left( \int_{\mathbb{R}} g_j^c (t) \right)} : g_j(t) = e^{-\lambda_j t^2}, \lambda_j > 0 \right\} = \inf \left\{ \frac{\det \left( \sum_{j=1}^{m} c_j \lambda_j u_j \otimes u_j \right)}{\prod_{j=1}^{m} \lambda_j^c} : \lambda_j > 0 \right\} = 1.$$

**Proof.** Let $g_j(t) = \exp(-\lambda_j t^2)$, $j = 1, \ldots, m$, where $\lambda_j$ are positive reals. Then,

$$\int_{\mathbb{R}^n} \prod_{j=1}^{m} g_j^c (\langle x, u_j \rangle) dx = \int_{\mathbb{R}^n} \exp \left( - \sum_{j=1}^{m} c_j \lambda_j \langle x, u_j \rangle^2 \right) dx$$

$$= \int_{\mathbb{R}^n} \exp \left( - \left( \sum_{j=1}^{m} c_j \lambda_j u_j \otimes u_j \right)(x), x \right) dx$$

$$= \frac{\pi^{n/2}}{\sqrt{\det \left( \sum_{j=1}^{m} c_j \lambda_j u_j \otimes u_j \right)}}.$$
On the other hand,

\[
\prod_{j=1}^{m} \left( \int_{\mathbb{R}} g_j \right)^{c_j} = \prod_{j=1}^{m} \left( \int_{\mathbb{R}} \exp(-\lambda_j t^2) dt \right)^{c_j} = \prod_{j=1}^{m} \left( \frac{\sqrt{\pi}}{\sqrt{\lambda_j}} \right)^{c_j} = \frac{\pi^{n/2}}{\sqrt{\prod_{j=1}^{m} \lambda_j^{c_j}}}
\]

since \( c_1 + \cdots + c_m = n \).

It follows that

\[
\inf \left\{ \left( \frac{\prod_{j=1}^{m} \left( \int_{\mathbb{R}} g_j \right)^{c_j}}{\int_{\mathbb{R}^n} \prod_{j=1}^{m} g_j^c \left( \langle x, u \rangle \right) dx} \right)^2 : g_j(t) = e^{-\lambda_j t^2}, \lambda_j > 0 \right\} = \inf \left\{ \frac{\det \left( \sum_{j=1}^{m} c_j \lambda_j u_j \otimes u_j \right)}{\prod_{j=1}^{m} \lambda_j^{c_j}} : \lambda_j > 0 \right\}.
\]

We turn to showing that this constant is 1.

By Lemma 7.5 we know that letting

\[
\lambda_I = \prod_{i \in I} \lambda_j \quad \text{and} \quad U_I = \left( \det \left( \sum_{j \in I} c_j u_j \otimes u_j \right) \right)^2,
\]

one has

\[
\sum_{|I|=n} U_I = 1.
\]

By the arithmetic-geometric means inequality,

\[
\sum_{|I|=n} \lambda_I U_I \geq \prod_{|I|=n} \lambda_I^{U_I} = \prod_{j=1}^{m} \lambda_j^{\sum_{(i,j \in I)} U_{I}}.
\]

Applying the Cauchy-Binet formula again (this time for all \( \lambda = 1 \) except the \( j^{th} \) which is 0, say), we have

\[
\sum_{\{I: j \in I\}} U_I = \sum_{|I|=n} U_I - \sum_{\{I: j \notin I\}} U_I = 1 - \det \left( \text{Id} - (\sqrt{c_j} u_j) \otimes (\sqrt{c_j} u_j) \right) = 1 - (1 - c_j |u_j|^2) = c_j.
\]

70
Going back to (20) and (22) we see that

\[
\det \left( \sum_{j=1}^{m} c_j \lambda_j u_j \otimes u_j \right) \geq \prod_{j=1}^{m} \lambda_j^{c_j}
\]

and thus

\[
\inf \left\{ \frac{\det \left( \sum_{j=1}^{m} c_j \lambda_j u_j \otimes u_j \right)}{\prod_{j=1}^{m} \lambda_j^{c_j}} : \lambda_j > 0 \right\} \geq 1.
\]

The choice \( \lambda_j = 1 \) gives equality in (23), which completes the proof.

7.3.3 The reverse inequality, and proof of both

We set

\[
I(f_1, \ldots, f_m) = \int_{\mathbb{R}^n} \prod_{j=1}^{m} f_j^{c_j}((x, u_j))dx.
\]

By considering the Gaussians, we have shown that,

\[
\sup \left\{ I(f_1, \ldots, f_m) : \int_{\mathbb{R}} f_j = 1, j = 1, \ldots, m \right\} \geq 1.
\]

The following is a reverse form of Theorem 7.2.

**Theorem 7.7** (Barthe). Let \( u_1, \ldots, u_m \in S^{n-1} \) and \( c_1, \ldots, c_m > 0 \) satisfy \( \text{Tr} = \sum_{j=1}^{m} c_j u_j \otimes u_j \).

If \( h_1, \ldots, h_m : \mathbb{R} \rightarrow \mathbb{R}^+ \) are measurable functions, we set

\[
K(h_1, \ldots, h_m) = \int_{\mathbb{R}^n} \sup \left\{ \prod_{j=1}^{m} h_j^{c_j}((\theta_j)) : \theta_j \in \mathbb{R}, x = \sum_{j=1}^{m} \theta_j c_j u_j \right\}dx.
\]

Then,

\[
\inf \left\{ K(h_1, \ldots, h_m) : \int_{\mathbb{R}} h_j = 1, j = 1, \ldots, m \right\} = 1.
\]

Before we continue, let us note that in dimension 1 this is Prékopa-Leindler. Indeed, we have a simple representation of identity by \( u_1 = e_1 \) and \( u_2 = -e_2 \) and \( c_1 = 1/2 \). Then the function integrated is \( \sup \{ f(y)^{1/2} g(z)^{1/2} : x = (y - z)/2 \} \). It is compared with the (square root of - if there was normalization of integrals equal to 1) product of the integrals.

Again as a first step we check what happens when we test the inequality on centered Gaussian functions. This will give us the easy part of this reverse Brascamp-Lieb inequality.

**Proposition 7.8.** With the notation of Theorem 7.7 we have

\[
\inf \left\{ K(h_1, \ldots, h_m) : \int_{\mathbb{R}} h_j = 1, j = 1, \ldots, m \right\} \leq 1.
\]

**Proof.** Let \( \lambda_j > 0, j = 1, \ldots, m \) and consider the functions \( h_j(t) = \exp(-t^2/\lambda_j) \). Then, the function

\[
m(x) := \sup \left\{ \prod_{j=1}^{m} h_j^{c_j}(\theta_j) : x = \sum_{j=1}^{m} \theta_j c_j u_j \right\}
\]

71
is given by

\[ m(x) = \exp \left( -\inf \left\{ \sum_{j=1}^{m} \frac{c_j \theta_j^2}{\lambda_j} : x = \sum_{j=1}^{m} \theta_j c_j u_j \right\} \right). \]

Note we claim that \( m \) is a quadratic form in \( x \) and in fact that it is given by \( \langle Bx, x \rangle \) where \( B = A^{-1} \) and \( A := \sum_{j=1}^{m} c_j \lambda_j u_j \otimes u_j \) (both are positive definite). To show this define

\[ \|x\|^2 = \sum_{j=1}^{m} c_j \lambda_j \langle x, u_j \rangle^2 = \langle Ax, x \rangle \]

and check that the dual norm is exactly

\[ \|y\|_*^2 = \inf \left\{ \sum_{j=1}^{m} \frac{c_j \theta_j^2}{\lambda_j} : y = \sum_{j=1}^{m} \theta_j c_j u_j \right\}. \]

You shall do this in the exercise sheet. Therefore,

\[ \|y\|^2_* = \langle By, y \rangle, \]

where \( B = A^{-1} \). It follows that

\[ \int_{\mathbb{R}^n} m(x) dx = \frac{\pi^{n/2}}{\sqrt{\det B}} = \pi^{n/2} \sqrt{\det A}. \]

On the other hand,

\[ \prod_{j=1}^{m} \left( \int_{\mathbb{R}} \exp(-t^2/\lambda_j) dt \right)^{c_j} = \pi^{n/2} \prod_{j=1}^{m} \lambda_j^{c_j/2}. \]

This shows that

\[ \inf \left\{ K^2(h_1, \ldots, h_m) : \int_{\mathbb{R}} h_j = 1 \right\} \]

\[ \leq \inf \left\{ \frac{\det \left( \sum_{j=1}^{m} c_j \lambda_j u_j \otimes u_j \right)}{\prod_{j=1}^{m} \lambda_j^{c_j/2}} : \lambda_j > 0 \right\} = 1 \]

and the proof is complete. \( \square \)

The main step in Barthe’s argument is the following proposition.

**Proposition 7.9.** Let \( f_1, \ldots, f_m : \mathbb{R} \to \mathbb{R}^+ \) and \( h_1, \ldots, h_m : \mathbb{R} \to \mathbb{R}^+ \) be integrable functions with

\[ \int_{\mathbb{R}} f_j(t) dt = \int_{\mathbb{R}} h_j(t) dt = 1, \quad j = 1, \ldots, m. \]

Then,

\[ I(f_1, \ldots, f_m) \leq K(h_1, \ldots, h_m). \]

**Proof.** We may assume that \( f_j, h_j \) are continuous and strictly positive. We use the transportation of measure idea that was used for the proof of the Prékopa-Leindler inequality: For every
We define $T_j : \mathbb{R} \to \mathbb{R}$ by the equation
\[
\int_{-\infty}^{T_j(t)} h_j(s) ds = \int_{-\infty}^{t} f_j(s) ds.
\]
Then, each $T_j$ is strictly increasing, 1-1 and onto, and
\[
T_j'(t) h_j(T_j(t)) = f_j(t), \quad t \in \mathbb{R}.
\]
We now define $W : \mathbb{R}^n \to \mathbb{R}^n$ by
\[
W(y) = \sum_{j=1}^{m} c_j T_j(\langle y, u_j \rangle) u_j.
\]
To compute the differential of this map we must differentiate $D(F(\langle x, u \rangle) u)$, and standard Calculus course tells us that this is $F'(\langle x, u \rangle) u \otimes u$. Thus
\[
D(W)(y) = \sum_{j=1}^{m} c_j T'_j(\langle y, u_j \rangle) u_j \otimes u_j.
\]
This implies
\[
\langle [D(W)(y)](v), v \rangle > 0 \quad \text{if} \quad v \neq 0
\]
and hence $W$ is in particular injective. Consider the function
\[
m(x) = \sup \left\{ \prod_{j=1}^{m} h_{cj}^{\theta_j} : x = \sum_{j=1}^{m} \theta_j c_j u_j \right\}.
\]
Then, (26) shows that
\[
m(W(y)) \geq \prod_{j=1}^{m} h_{cj}^{\theta_j}(T_j(\langle y, u_j \rangle))
\]
for every $y \in \mathbb{R}^n$. It follows that
\[
\int_{\mathbb{R}^n} m(x) dx \geq \int_{W(\mathbb{R}^n)} m(x) dx
\]
\[
= \int_{\mathbb{R}^n} m(W(y)) \cdot | \det D(W)(y) | dy
\]
\[
\geq \int_{\mathbb{R}^n} \prod_{j=1}^{m} h_{cj}^{\theta_j}(T_j(\langle y, u_j \rangle)) \det \left( \sum_{j=1}^{m} c_j T'_j(\langle y, u_j \rangle) u_j \otimes u_j \right) dy
\]
\[
\geq \int_{\mathbb{R}^n} \prod_{j=1}^{m} h_{cj}^{\theta_j}(T_j(\langle y, u_j \rangle)) \prod_{j=1}^{m} (T'_j(\langle y, u_j \rangle))^{c_j} dy,
\]
where in the last inequality we have used Proposition 7.6. Therefore, taking (25) into account we have
\[
\int_{\mathbb{R}^n} m(x) dx \geq \int_{\mathbb{R}^n} \prod_{j=1}^{m} f_j^{c_j}(\langle y, u_j \rangle) dy = I(f_1, \ldots, f_m).
\]
In other words, \( I(f_1, \ldots, f_m) \leq K(h_1, \ldots, h_m). \)

\[ \begin{align*} 
\int_{\mathbb{R}} f_j(t)dt &= \int_{\mathbb{R}} h_j(t)dt = 1, \quad j = 1, \ldots, m. 
\end{align*} \]

Then,

\[ I(f_1, \ldots, f_m) \leq K(h_1, \ldots, h_m). \]

Taking the supremum over all such functions \( f_i \) and the infimum over all such functions \( h_i \) we get that

\[ 1 \leq \sup \left\{ I(f_1, \ldots, f_m) \right\} \leq \inf \left\{ K(h_1, \ldots, h_m) \right\} \leq 1, \]

so that there must be equality throughout.

\[ \square \]

8 Some remarks on Measure Concentration

Concentration of measure is a phenomenon of high dimensions which is responsible to many counter-intuitive (until intuition changes and the become intuitive) results. Let me approach it from a very non-standard angle:

Lemma 8.1 (Cauchy formula). Let \( K \in \mathbb{K}^n. \) Then

\[ \text{Vol}_{n-1}(\partial K) = \frac{n \kappa_n}{\kappa_{n-1}} \int_{S^{n-1}} \text{Vol}_{n-1}(P_{u \perp} K) d\sigma(u). \]

Sometimes one uses integration on the sphere with respect to usual Lebesgue measure, the relation being \( d\sigma(u) = du/\text{Vol}_{n-1}(S^{n-1}) = du/(n\kappa_n) \) so we in these terms you get the formula

\[ \text{Vol}_{n-1}(\partial K) = \frac{1}{\kappa_{n-1}} \int_{S^{n-1}} \text{Vol}_{n-1}(P_{u \perp} K) du \]

which may be more familiar to you as “Cauchy’s formula for surface area”.

Proof. We work with a polytope \( P \) first. For a generic \( u \in S^{n-1} \), each facet \( F_i \) of \( P \) has some angle \( \theta_i \) between its normal and \( u \) which is not \( \pi/2 \). When such a facet is projected onto \( u \perp \), its area when projected is \( \lvert \cos(\theta) \rvert \) times its original area. Clearly the projection \( P_{E}(P) \) is covered twice by the projections of the facets of \( P \). We get that

\[ \text{Vol}_{n-1}(P_{u \perp}(P)) = \frac{1}{2} \sum_{i=1}^{m} \lvert \cos(\theta_i) \rvert \text{Vol}_{n-1}(P_i). \]

When integrated over the sphere, we get

\[ \int_{S^{n-1}} \text{Vol}_{n-1}(P_{u \perp}(P)) d\sigma(u) = \frac{1}{2} \sum_{i=1}^{m} \text{Vol}_{n-1}(P_i) \int_{S^{n-1}} \lvert \cos(\theta(n_i, u)) \rvert d\sigma(u). \]

By rotation invariance, the latter integral does not depend on \( n_i \) and in simply a constant depending on the dimension (for example, let \( n = (1, 0, \ldots, 0) \), so that \( \cos(\theta) = u_1 \)). We have
thus shown that there exists some constant $c_n$ such that for any polytope $P$ we have

$$\int_{S^{n-1}} \text{Vol}_{n-1} P_u^\perp (P) d\sigma(u) = c_n \text{Vol}_{n-1}(\partial P).$$

Since both sides are well defined for convex bodies and are clearly monotone, and since we have equality for all polytopes, we have equality for bodies as well. Thus, for the same $c_n$ (defined for now by half the integral of $|u_1|$ over $u \in S^{n-1}$) we get for all convex $K$ that

$$\int_{S^{n-1}} \text{Vol}_{n-1} P_u^\perp (K) d\sigma(u) = c_n S_n(K).$$

Instead of computing the integral, we may simply plug in $K = B_2^n$. \hfill $\square$

Looking at the proof above, we see that (along with proving a nice formula for surface area) we computed

$$\int_{S^{n-1}} |x_1| d\sigma(x) = \frac{2}{\text{Vol}_{n-1}(\partial K)} \int_{S^{n-1}} \text{Vol}_{n-1}(P_u K) d\sigma(u) = \frac{2\kappa_{n-1}}{n\kappa_n}.$$

Since we know what is the asymptotic of $\kappa_n$ we may use it to compute:

$$\frac{2\kappa_{n-1}}{n\kappa_n} = \frac{2}{n} \frac{\pi^{\frac{n-1}{2}} \Gamma\left(\frac{n}{2} + 1\right)}{\Gamma\left(\frac{n}{2} + 1\right)} \approx \frac{2}{n\sqrt{\pi}} \frac{(n + 1)}{2}^{1/2} \approx \sqrt{2/\pi} \frac{1}{\sqrt{n}}.$$

So, the average of $|u_1|$ on the sphere is of the order $n^{-1/2}$, which is quite small. If you think of “median” instead of average, and up to factor 2 the median of a positive quantity is less than the average, this means that about $1/2$ of the measure of the whole sphere is concentrated near the hyperplane $x_1 = 0$, in a strip of width $1/\sqrt{n}$.

Concentration is not just about half the volume, but about the majority of volume. Indeed, one may compute by integrating $\sigma\{x \in S^{n-1} : |x_1| < r\}$, to see how it behaves with respect to $r$. The above computation show that this integral will be $1/2$ when $r$ is of the order $n^{-1/2}$. One may write out this integral (it is just powers of cos) and estimate it. However, to avoid these computations, we can estimate it pretty well (say for $r < 1/2$) by the following trick: Draw a picture of the ball, and a point $x$ at distance $r$ from its center, say in direction $e_1$. Then the cap $C_r = \{x \in S^{n-1} : |x_1| \geq r\}$ is a subset of a ball centered at $x = re_1$ of radius $\sqrt{1 - r^2}$. In fact, not only the cap but also its convex hull with the origin, $A = \text{conv}(C_r, \{0\})$, lies within this ball. Therefore

$$\sigma\{u : u_1 > r\} = \frac{\text{Vol}_{n}(A)}{\text{Vol}_{n}(B_2^n)} \leq \frac{\text{Vol}_{n}((1 - r^2)^{1/2}B_2^n)}{\text{Vol}_{n}(B_2^n)} = (1 - r^2)^{n/2} \leq e^{-\frac{1}{2}nr^2}.$$

This is not an exact inequality, but it suffices for most applications.

We add this observation to the isoperimetric inequality on the sphere (we did not prove it, but it can be proved using Steiner symmetrization on the sphere, and also in other ways) which says that of all sets of given measure on $S^{n-1}$, the one which has the smallest $\varepsilon$-extension is a cap. We use it below for measure 1/2, in which case the cap is a half-sphere. We conclude:

**Proposition 8.2.** For any $A \subset S^{n-1}$ with $\sigma(A) = 1/2$, and any $\varepsilon < 1/2$ we have

$$\sigma(A_\varepsilon) = \sigma\{u \in S^{n-1} : d(u, A) < \varepsilon\} \geq 1 - e^{-\varepsilon^2 n}.$$
I did not specify which \( d \), is it euclidean distance or geodesic, and instead of \( 1/2 \) I put \( c \), so that this unspecification is of no consequence.

**Remark 8.3.** One should compare Proposition 8.2 with Borell’s Lemma, namely Theorem 5.5 from Section 5.2:

Let \( K \) be a convex body in \( \mathbb{R}^n \) with volume \( \text{Vol}_n(K) = 1 \), and let \( A \) be a closed, convex and centrally symmetric set such that \( \text{Vol}_n(K \cap A) = \delta > \frac{1}{2} \). Then, for every \( t > 1 \) we have \( \text{Vol}_n(K \cap (tA)^c) \leq \delta \left( \frac{1 - \delta}{\delta} \right)^{\frac{n+1}{2}}. \)

Which is a concentration result as well, but of a somewhat different nature.

**Corollary 8.4.** For any continuous \( f : S^{n-1} \rightarrow \mathbb{R} \) which is \( L \)-lipshitz and any \( \delta < L/2 \) we have for \( M \) the median of \( f \)

\[
\sigma \{ u \in S^{n-1} : |f(u) - M| < \delta \} \geq 1 - 2 \exp^{-c \delta^2 n / L^2}.
\]

**Proof.** Let \( M \) be the median of \( f \) so that \( \sigma(A_1 = \{ u \in S^{n-1} : f(x) \leq M \}) \geq 1/2 \) as is \( \sigma(A_2 = \{ u \in S^{n-1} : f(x) \geq M \}) \). Then

\[
\sigma \{ u \in S^{n-1} : d(u, A_1) < \varepsilon \text{ and } d(u, A_2) \} \geq 1 - 2 \exp^{-c \varepsilon^2 n}.
\]

Therefore

\[
\sigma \{ u \in S^{n-1} : |f(u) - M| < \varepsilon L \} \geq 1 - 2 \exp^{-c \varepsilon^2 n}.
\]

We can reformulate as

\[
\sigma \{ u \in S^{n-1} : |f(u) - M| < \varepsilon M \} \geq 1 - 2 \exp^{-c (\varepsilon M / L)^2 n}.
\]

where we see that what matters in this estimate is the ratio \( M/L \) somehow.

We plan to show how concentration can help prove theorems, and we shall examplify this using Dvoretzky’s theorem about near-Euclidean sections of a convex body. For lack of time we do not go into the history of the theorem. We shall show Milman’s proof. The sketch is as follows: One is given a centrally symmetric convex body in dimension \( n \), and would like to find a section by a subspace \( E \) of dimension \( \dim(E) = k \) such that \( K \cap E \) is almost an ellipsoid, up to some \( \varepsilon \). That is, to find a Euclidean norm \( |\cdot| \) such that for \( x \in E \) we have

\[
(1 - \varepsilon)|x| \leq \|x\|_K \leq (1 + \varepsilon)|x|.
\]

To this end, one takes a net on the sphere of an \( k \)-dimensional space, and uses concentration phenomenon to show that for some rotation \( U \in O(n) \), all the elements \( U(x_i) \) for \( x_i \) in the net, have the same norm more or less. Then to pass from a net to the whole sphere is not so difficult (it is called successive approximation).

One then checks the paremeters: for the whole net to work, it needs to have cardinality to exceeding the one allowed by concentration. This tells you what is the maximal dimension \( n \) you are allowed to take.

In the concentration part, we saw that what mattered was \( M/L \), the smaller it was, we were in bigger trouble. Here is where John’s position enters. We are allowed, to begin with, to change the body linearly. A euclidean section will then become an ellipsoidal one, which is again euclidean. So we may assume without loss of generality that we are in John position.
In this case, the euclidean norm we find will be the standard one. So we must explain why in John position, the parameter $M/L$ is not too small. This will be our first step.

We will use the following corollary from Dvoretzky Rogers Lemma which you showed in the exercises:

Assume that $B^n_2$ is the maximal volume ellipsoid of a convex body $K$. We can find orthonormal vectors $z_1, \ldots, z_k$ such that $\frac{1}{4} \leq \|z_j\| \leq 1$ for all $j = 1, \ldots, n$.

We define, for the function $r(x) = \|x\|$, three parameters, its mean and its median, which should be thought of as very close, and its maximum (on the sphere), which is also its Lipschitz constant.

**Definition 8.5.** Let $X = (\mathbb{R}^n, \| \cdot \|)$ have unit ball $K$. We define

$$M = M(X_K) = \int_{S^{n-1}} \|x\| d\sigma(x),$$

We set $L_r$ the median of $r$ on $S^{n-1}$, and $b = \max_{x \in S^{n-1}} \|x\|$, so that $\|x\| \leq b|x|$ for all $x \in \mathbb{R}^n$.

Note that by the definition of $b$ we have that for some unit vector $u \in S^{n-1}$ we have that $K \subseteq S = \{ : \langle x, u \rangle \leq 1/b \}$, so

$$M(K) \geq M(S) = \int_{S^{n-1}} \|x\| d\sigma(x) = b \int_{S^{n-1}} |x_1| d\sigma(x) = b/\sqrt{n},$$

which gives a lower bound, valid in general, for $M/b$. Similarly $L_r$ is monotone so

$$L_r(K) \geq L_r(S) = \{ t : \sigma(x : \|x\| > t) = 1/2 \} \simeq cb/\sqrt{n}.$$

Next let us explain why $L_r$ and $M$ are in fact equivalent:

**Lemma 8.6.** There exists a universal constant such that for any $K = -K$ convex body we have

$$\frac{1}{2} \leq \frac{M}{L_r} \leq c,$$

Moreover, we have the estimate which is sometimes more useful

$$\left| \frac{M}{L_r} - 1 \right| \leq c' \frac{b}{\sqrt{nL}},$$

for universal $c'$.

**Proof.** First observe that

$$M = \int_{S^{n-1}} \|x\| d\sigma(x) \geq \int_{\{x : \|x\| \geq L_r\}} \|x\| d\sigma(x) \geq \frac{1}{2} L_r.$$ 

Therefore, $M/L_r \geq 1/2$.

For the other direction, and the second inequality: We know that for every $\varepsilon > 0$

$$\sigma(\{ x : |r(x) - L_r| \geq b\varepsilon \}) \leq 2\exp(-c_1 \varepsilon^2 n).$$
We write
\[ |M - L_r| \leq \int_{S^{n-1}} |r(x) - L_r| \, d\sigma(x) = \int_0^\infty \sigma\{x : |r(x) - L_r| \geq t\} \, dt.\]

Set \( b \varepsilon = t \) to get
\[ |M - L_r| \leq \int_0^\infty 2\exp\left(-c_1 t^2 \frac{n}{b^2}\right) \, dt = c_2 \frac{b}{\sqrt{n}}.\]

Therefore,
\[ \left| \frac{M}{L_r} - 1 \right| \leq c_3 \frac{b}{\sqrt{n}L_r},\]
so in particular
\[ M \leq L_r + c_3 \frac{b}{\sqrt{n}} \leq L_r(1 + c_4)\]
where we used the fact, proven above, that \( L_r \geq cb/\sqrt{n} \).

We will thus estimate from below the number \( M/b \) in John’s position, rather than \( L_r/b \). To this end we will need to know an estimate for one more integral of a simple function on \( S^{n-1} \):

**Lemma 8.7.** For every \( 1 \leq m \leq n \), denoting \( x = (x_1, \ldots, x_n) \), we have that
\[
\int_{S^{n-1}} \max_{1 \leq j \leq m} |x_j| \, d\sigma(x) \geq c_4 \left( \frac{\log m}{n} \right)^{1/2},
\]
where \( c_4 > 0 \) is an absolute constant.

**Proof.** the trick here is to move to gaussian random variables instead of spherical. Integration in polar coordinates shows that (for any 1-homogeneous function)
\[
\int_{\mathbb{R}^n} \max_{1 \leq j \leq m} |t_j| \, d\gamma_n(t) = \lambda_n \int_{S^{n-1}} \max_{1 \leq j \leq m} |x_j| \, d\sigma(x),
\]
where \( \lambda_n \simeq \sqrt{n} \). However,
\[
\int_{\mathbb{R}^n} \max_{1 \leq j \leq m} |t_j| \, d\gamma_n(t) = \int_{\mathbb{R}^m} \max_{1 \leq j \leq m} |t_j| \, d\gamma_m(t).
\]
So we would like to estimate the latter integral. Using the inequality
\[
\frac{1}{\sqrt{2\pi}} \int_s^\infty e^{-t^2/2} \, dt \geq \frac{1}{\sqrt{2\pi}} \frac{s}{s^2 + 1} e^{-s^2/2}
\]
which is valid for all \( s > 0 \), we may write
\[
\gamma_m\left(\left\{ t : \max_{1 \leq j \leq m} |t_j| < s \right\}\right) = (2\pi)^{-m/2} \int_{-s}^s \cdots \int_{-s}^s \exp\left(-\frac{1}{2} \sum_{j=1}^m t_j^2\right) \, dt_1 \cdots dt_m
\]
\[
= \left( \frac{1}{\sqrt{2\pi}} \int_{-s}^s e^{-t^2/2} \, dt \right)^m
\]
\[
\leq \left( 1 - \frac{2s}{(s^2 + 1)\sqrt{2\pi}} e^{-s^2/2} \right)^m.
\]
If we choose \( s = \sqrt{\log m} \), (assume \( m \) is large enough), we see that
\[
\gamma_m \left( \left\{ t : \max_{j \leq m} |t_j| \geq \sqrt{\log m} \right\} \right) \geq \frac{1}{2}.
\]
Then,
\[
\int_{S^{n-1}} \max_{j \leq m} |x_j| d\sigma(x) \approx \frac{1}{\sqrt{n}} \int_{\mathbb{R}^m} \max_{j \leq m} |t_j| d\gamma_m(t) \geq \frac{\sqrt{\log m}}{\sqrt{n}} \gamma_m \left( \left\{ t : \max_{j \leq m} |t_j| \geq \sqrt{\log m} \right\} \right) \geq \frac{1}{2} \left( \frac{\log m}{n} \right)^{1/2},
\]
and the result follows.

Using this estimate, we are now able to prove that in John’s position \( M/b \) is well bounded from below.

**Theorem 8.8.** Let \( X = (\mathbb{R}^n, \| \cdot \|) \) and assume that \( B^n_2 \) is the maximal volume ellipsoid of its unit ball \( K \). Then
\[
\frac{M}{b} \geq c \sqrt{\frac{\log n}{n}}.
\]

**Proof.** By the Dvoretzky-Rogers lemma we can find an orthonormal basis \( \{ x_1, \ldots, x_n \} \) with \( \|x_i\| \geq 1/4 \) for all \( i = 1, \ldots, n \). Note that
\[
M = \int_{S^{n-1}} \left\| \sum_{i=1}^n a_i x_i \right\| d\sigma(a) = \int_{S^{n-1}} \int_{\mathbb{E}^n_2} \left\| \sum_{i=1}^n \varepsilon_i a_i x_i \right\| d\varepsilon d\sigma(a).
\]
We shall use the fact that for all \( j = 1, \ldots, n \),
\[
\int_{\mathbb{E}^n_2} \left\| \sum_{i=1}^n \varepsilon_i y_i \right\| d\varepsilon \geq \|y_j\|,
\]
which can be proved using the triangle inequality and induction, see Exercise sheet 11. Setting \( y_i = a_i x_i \) we have
\[
\int_{\mathbb{E}^n_2} \left\| \sum_{i=1}^n \varepsilon_i a_i x_i \right\| d\varepsilon \geq \max_{1 \leq i \leq n} \|a_i x_i\|.
\]
Thus, we get
\[
M \geq \int_{S^{n-1}} \max_{1 \leq i \leq n} \|a_i x_i\| d\sigma(a).
\]
Since \( \|x_i\| \geq 1/4 \), we get
\[
M \geq \frac{1}{4} \int_{S^{n-1}} \max_{1 \leq i \leq n} |a_i| d\sigma(a),
\]
and Lemma 8.7 gives
\[
M \geq c \sqrt{\frac{\log n}{n}}.
\]
The proof is complete. \( \square \)
Let us look at the dimensions for a minute: In the John position, using the bound we just obtained, the probability estimate (on $S^{n-1}$) for $||x|| - L_r > \delta L_r$ is $\exp(-c\delta^2 n(M/L_r)^2) \leq \exp(-c\log n\delta^2)$. Since we shall see below, we cannot expect $\delta$’s help (though it will be a constant, so not interfering too much) this means we can take care only on a power of $n$ points; which should (for the net to exist, as we shall see below) needs to be $\exp(k)$ where $k$ is the dimension of the section. So, $k = \log n$ is the right order of dimension. We now do this formally.

To use concentration of measure in order to find a full subspace, we use a $\delta$-net on the sphere of a $k$-dimensional subspace, namely a set $\mathcal{N} \subset S^{k-1}$ such that for every $y \in S^{k-1}$ there exists $x \in \mathcal{N}$ such that $|x - y| < \delta$. We use the following lemma

**Lemma 8.9.** Let $\delta > 0$ and $k \in \mathbb{N}$. There is a $\delta$-net $\mathcal{N}$ for $S^{k-1}$ with cardinality $|\mathcal{N}| \leq (1 + \frac{2}{\delta})^k$.

**Proof.** Let $\{x_i\}_{i=1}^N$ be a maximal $\delta$-separated set in $S^{k-1}$, namely such that any two of its points have distance greater than or equal to $\delta$. Then $\{x_i\}_{i=1}^N$ is a $\delta$-net for $S^{k-1}$. Since the sets $x_i + \frac{1}{2}\delta B_2^k$ have disjoint interiors and are all included in $B_2^k + \frac{\delta}{2} B_2^k$, we get that

$$N(\delta) \Vol_k(B_2^k) \leq (1 + \frac{\delta}{2})^k \Vol_k(B_2^k),$$

which gives $N \leq (1 + \frac{2}{\delta})^k$.

Let $X = (\mathbb{R}^n, \|\cdot\|)$ be an $n$-dimensional normed space. Recall that the function $r : S^{n-1} \to \mathbb{R}$ is defined by $r(x) = \|x\|$ and is Lipschitz continuous with constant $b$ which is the smallest positive real number such that $\|x\| \leq b|x|$ for all $x \in \mathbb{R}^n$.

**Lemma 8.10.** Let $X = (\mathbb{R}^n, \|\cdot\|)$ be an $n$-dimensional normed space with $\|x\| \leq b|x|$ for all $x \in X$. Assume $m \leq \frac{1}{4} \exp(c_1 \varepsilon^2 n/2)$ and $y_i \in S^{n-1}, i = 1, \ldots, m$. There exists $B \subset O(n)$ with $\nu_n(B) \geq 1 - \exp(-c_1 \varepsilon^2 n/2)$ such that

$$L_r - b\varepsilon \leq \|Uy_i\| \leq L_r + b\varepsilon$$

for every $U \in B$ and $i = 1, \ldots, m$.

**Proof.** If $x_0 \in S^{n-1}$ then, for every $A \subset S^{n-1},$

$$\sigma(A) = \nu_n(\{U \in O(n) : Ux_0 \in A\}),$$

where $\nu$ is the Haar measure on $O(n)$. Consider the set

$$A = \{x \in S^{n-1} : L_r - b\varepsilon \leq \|x\| \leq L_r + b\varepsilon\}.$$

By concentration

$$\sigma(A) \geq 1 - 4e^{-c_1 \varepsilon^2 n}.$$

It follows that if we define

$$B_i = \{U \in O(n) : L_r - b\varepsilon \leq \|Uy_i\| \leq L_r + b\varepsilon\}$$

for $i = 1, \ldots, m$, then

$$\nu_n(B_i) > 1 - 4e^{-c_1 \varepsilon^2 n}.$$
Proposition 8.12. Let \( \delta \) be a subspace \( F \).

Proof. Let \( \delta \) be an \( n \)-dimensional normed space with \( \| x \| \leq b \| x \| \) for all \( x \in X \). Let \( \delta, \varepsilon \in (0, 1) \). If \( (1 + 2/\delta)^k \leq \frac{1}{4} \exp(c_1 \varepsilon^2 n/2) \) then we may find \( \Gamma \subset G_{n,k} \), with \( \nu_{n,k}(\Gamma) \geq 1 - \exp(-c_1 \varepsilon^2 n/2) \), such that: for every \( F \in \Gamma \) there exists a \( \delta \)-net \( \mathcal{N}_F \) for \( S_F \) such that

\[
L_r - b \varepsilon \leq \| x \| \leq L_r + b \varepsilon
\]

for all \( x \in \mathcal{N}_F \).

Proof. Let \( \delta, \varepsilon \in (0, 1) \) and \( k \in \mathbb{N} \) satisfy the condition in the statement of the proposition. We fix a subspace \( F_0 \) of \( \mathbb{R}^n \) with \( \text{dim}(F_0) = k \). From Lemma 8.9 we can find a \( \delta \)-net \( \{ y_1, \ldots, y_m \} \) for the unit sphere \( S_{F_0} = S^{n-1} \cap F_0 \) of \( F_0 \), with \( m \leq (1 + 2/\delta)^k \). Consider the set \( B \subset O(n) \) which is promised by Lemma 8.10. If \( U \in B \) and \( 1 \leq i \leq m \) then

\[
M - b \varepsilon \leq \| U y_i \| \leq M + b \varepsilon.
\]

We set \( F_U := U(F_0) \) and \( x_i := U y_i \) for \( i = 1, \ldots, m \). Since \( U \) is an orthogonal transformation, \( \{ x_1, \ldots, x_m \} \) is a \( \delta \)-net for \( S_{F_U} \), which satisfies

\[
M - b \varepsilon \leq \| x_i \| \leq M - b \varepsilon.
\]

We set \( \Gamma = \{ F_U : U \in B \} \) and observe that

\[
\nu_{n,k}(\Gamma) = \nu_{n,k}(\{ U(F_0) : U \in B \}) = \nu_n(B).
\]

Using now the fact that \( \| \cdot \| \) is a norm, we will extend, for \( F \in \Gamma \), the estimate of Proposition 8.11 from a \( \delta \)-net \( \mathcal{N}_F \) of \( S_F \) to the whole sphere \( S_F \), by the method of successive approximation.

Proposition 8.12. Let \( X = (\mathbb{R}^n, \| \cdot \|) \) be an \( n \)-dimensional normed space with \( \| x \| \leq b \| x \| \) for all \( x \in X \). Let \( \delta, \varepsilon \in (0, 1) \). If \( (1 + 2/\delta)^k \leq \frac{1}{4} \exp(c_1 \varepsilon^2 n/2) \) then we may find \( \Gamma \subset G_{n,k} \), with \( \nu_{n,k}(\Gamma) \geq 1 - \exp(-c_1 \varepsilon^2 n/2) \), such that: for every \( F \in \Gamma \) and for every \( y \in S_F \) we have

\[
(27) \quad \frac{1 - 2 \delta}{1 - \delta} M - \frac{b \varepsilon}{1 - \delta} \leq \| y \| \leq \frac{M + b \varepsilon}{1 - \delta}.
\]

Proof. The subset \( \Gamma \) will be the same as given by Proposition 8.11. Let \( y \in S_F \). There exists \( x_0 \in \mathcal{N}_F \) such that \( |y - x_0| = \delta_1 < \delta \). Then, \( \frac{y - x_0}{\delta_1} \in S_F \), and hence we can find \( x_1 \in \mathcal{N}_F \) with
\[ \frac{|y-x_0|}{\delta} - x_1 = \delta_2 < \delta. \] Then,
\[ |y - x_0 - \delta_1 x_1| = \delta_1 \delta_2 < \delta^2. \]
Inductively, we find \( x_0, \ldots, x_n \in \mathcal{N}_F \) and \( \delta_1, \ldots, \delta_n \) such that
\[ |y - \sum_{i=0}^{n} (\prod_{j=0}^{i} \delta_j) x_i| \leq \delta^{n+1}, \]
where \( \delta_0 = 1, \delta_i < \delta \) and so \( \prod_{j=0}^{i} \delta_j \leq \delta^i \). Since \( \delta < 1 \),
\[ y = \sum_{i=0}^{\infty} (\prod_{j=0}^{i} \delta_j) x_i. \]
We get
\[ \|y\| = \left\| \sum_{i=0}^{\infty} \left( \prod_{j=0}^{i} \delta_j \right) x_i \right\| \leq \sum_{i=0}^{\infty} \delta^i \|x_i\| \leq (M + b\varepsilon) \sum_{i=0}^{\infty} \delta^i = \frac{M + b\varepsilon}{1 - \delta}. \]
On the other hand,
\[ \|y\| \geq \|x_0\| - \left\| \sum_{i=1}^{\infty} \left( \prod_{j=0}^{i} \delta_j \right) x_i \right\| \geq M - b\varepsilon - \frac{\delta}{1 - \delta} (M + b\varepsilon) = \frac{1 - 2\varepsilon}{1 - \delta} M - \frac{b\varepsilon}{1 - \delta}. \]
It follows that (27) holds true for all \( y \in S_F \).

Finally, all that is left is to put all the ingredients together to form

**Theorem 8.13** (Milman). Let \( X = (\mathbb{R}^n, \| \cdot \|) \) and assume that \( \|x\| \leq b|x| \) for all \( x \in \mathbb{R}^n \). For any \( \varepsilon \in (0, 1) \) if
\[ k \leq k_X(\varepsilon) = c_2 \varepsilon^2 \log^{-1} \left( \frac{2}{\varepsilon} \right) n(M/b)^2, \]
then we can find a subspace \( F \) of \( \mathbb{R}^n \) with \( \dim(F) = k \) such that for every \( x \in S_F \)
\[ (1 - \varepsilon) L_r \leq \|x\| \leq (1 + \varepsilon) L_r. \]
Moreover, this holds for all \( F \) in a subset \( A \subset G_{n,k} \) of measure at least \( \nu_{n,k}(A) \geq 1 - \exp(-c_3 \varepsilon^2 k) \). In particular, if \( K \) is in John position, one can find an \( \varepsilon \)-euclidean subspace for all \( k \leq c(\varepsilon) \log n \).

Finally, let us explain why the \( \log(n) \) term cannot be improved (the dependence on \( \varepsilon \) can, actually).

**Proposition 8.14.** The cube \([-1, 1]^n\) has \( \varepsilon \)-euclidean section of dimension \( c(\varepsilon) \log n \) and not more.

The proof of the proposition will follow directly from the following lemma (applied to the dimension \( k \simeq \log n \)), as sections of a cube \([-1, 1]^n\) have at most \( 2n \) facets. Note, further (it’s not needed for the proof but good to note), that any centrally symmetric polytope in \( \mathbb{R}^k \) with \( m \) pairs of opposite facets is in fact (up to a linear transformation) a section of a cube in \( \mathbb{R}^m \).
Indeed if for some family of vectors $u_i \in \mathbb{R}^k$ we have

$$P = \{ x \in \mathbb{R}^k : |\langle x, u_i \rangle| \leq 1 \}$$

then using the linear map $T$ given by

$$x \mapsto Tx = (\langle x, u_1 \rangle, \ldots, \langle x, u_m \rangle) \in \mathbb{R}^m$$

and letting $T\mathbb{R}^k := E_k \subset \mathbb{R}^m$ a $k$-dimensional subspace, then

$$TP = E_k \cap [-1, 1]^n.$$

**Lemma 8.15.** Let $P$ be a polytope in $\mathbb{R}^k$ which is $R$-close in geometric distance to the euclidean ball, that is, $aB_2^k \subset K \subset aRB_2^k$. Then the number of facets of $P$ is at least $\exp(k/2R^2)$.

**Proof.** We may without loss of generality let $a = 1$ otherwise we rescale $P$, the number of facets is not affected. Assume that $P$ is given by $P = \{ x \in \mathbb{R}^k : \langle x, u_i \rangle \leq 1 \}$ (we may have assumed central symmetry, since it only helps, but it is of no importance here). Since $B_2^n \subset P$ clearly $u_i/|u_i| \in B_2^n \subset P$ and so $\langle u_i, u_i \rangle \leq 1$ and so $u_i \in B_2^n$ (it is likely that if $u_i$’s are optimized then they will actually be in $S^{n-1}$ but we will not assume this). Since $P \subset RB_2^n$ we know that for any point $u \in S^{n-1}$, the point $Ru$ is not in the interior of $P$, so that for at least one vector $u_i$ we must have $\langle u, u_i \rangle \geq R^{-1}$. Therefore, the whole sphere $S^{n-1}$ is covered by the caps $C_i = \{ u : \langle u, u_i \rangle \geq R^{-1} \}$. As is the previous lesson, we know that the measure, on $S^{n-1}$, of such a cap is very small. More precisely, at least when $R > 2$, we know $\sigma(C_i) \leq \exp(-n/(2R^2))$, which completes the proof. \qed