Convex Geometric Analysis

Contents

Chapter 1

Brunn-Minkowski inequality

1.1 Introduction

The Brunn-Minkowski inequality gives a fundamental relation between Minkowski addition and volume in $\mathbb{R}^n$.

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

$$|K+T|^{1/n} \geq |K|^{1/n} + |T|^{1/n}. \tag{1.1}$$

If we make the additional hypothesis that K and T are convex bodies, then we can have equality in (1.1) only if K and T are homothetical.

The inequality expresses in a sense the fact that volume is a **concave** function with respect to Minkowski addition. For this reason, it is often written in the following form: If K, T are non-empty compact subsets of $\mathbb{R}^n$ and $\lambda \in (0,1)$, then

$$|\lambda K + (1-\lambda)T|^{1/n} \geq \lambda |K|^{1/n} + (1-\lambda)|T|^{1/n}. \tag{1.2}$$

Using (1.2) and the arithmetic-geometric means inequality we can also write

$$|\lambda K + (1-\lambda)T| \geq |K|^{\lambda}|T|^{1-\lambda}. \tag{1.3}$$

This weaker form of the Brunn-Minkowski inequality has the advantage of being dimension free. It is actually equivalent to (1.1) in the following sense: if we know that (1.3) holds for all K, T and λ, we can recover (1.1) as follows.

Consider non-empty compact sets K and T (we may assume that $|K| > 0$ and $|T| > 0$; otherwise there is nothing to prove), and define

$$K_1 = |K|^{-1/n}K \quad , \quad T_1 = |T|^{-1/n}T \quad , \quad \lambda = \frac{|K|^{1/n}}{|K|^{1/n} + |T|^{1/n}}.$$

Then, K_1 and T_1 have volume 1, and hence (1.3) gives

$$|\lambda K_1 + (1-\lambda)T_1| \geq 1. \tag{1.4}$$

Since

$$\lambda K_1 + (1-\lambda)T_1 = \frac{K+T}{|K|^{1/n} + |T|^{1/n}},$$

we immediately get (1.1).

There are many interesting proofs of the Brunn-Minkowski inequality, all of them related to important ideas. In the next three sections we discuss a variety of arguments which imply (1.1).

In Section 1.2 we present some classical proofs of the Brunn-Minkowski inequality. The first one is the original proof of Minkowski which is based on the Brunn concavity principle. We give two proofs of Brunn's theorem: one goes by Schwarz symmetrization, while the other one uses a bisection argument of Gromov and Milman. The second proof works for (not necessarily convex) compact sets: it is based on elementary sets and is due to Hadwiger and Ohmann. The third proof clarifies the equality case for convex bodies: it is due to Kneser and Süss and goes by induction on the dimension. Finally, we prove Bonnesen's refinements of the Brunn-Minkowski inequality: a linear inequality for convex bodies whose projections onto some hyperplane have the same $(n-1)$-dimensional volume and another version involving the volumes of hyperplane sections.

In Section 1.3 we introduce two volume preserving, one to one and onto transformations between two open convex bodies: the Knöthe map and the Brenier map. Both maps can be used to give proofs of the Brunn-Minkowski inequality, but they will also be quite useful tools for several results in subsequent chapters.

In Section 1.4 we prove the Prékopa-Leindler inequality, which may be viewed as a functional form of the Brunn-Minkowski inequality. As an application, we prove the logarithmic Sobolev inequality and discuss some of its consequences in Gauss space.

Finally, in Section 1.5 we give a first list of important geometric applications of (1.1) and its relatives. This is far from being complete but shows the fundamental character and strength of the Brunn-Minkowski inequality.

1.2 Classical proofs

1.2.1 Brunn's concavity principle

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

Theorem 1.2.1 *Let K be a convex body in $\mathbb{R}^n$ and let F be a k-dimensional subspace. Then, the function $f : F^{\perp} \to \mathbb{R}$ defined by $f(x) = |K \cap (F+x)|^{1/k}$ is concave on its support.*

Sketch of the proof: The proof goes by symmetrization. The **Steiner symmetrization** of K in the direction of $\theta \in S^{n-1}$ is the set $S_\theta(K)$ consisting of all points of the form $x + \lambda\theta$, where x is in the projection $P_{\theta^\perp}(K)$ of K onto $\theta^\perp$ and $|\lambda| \leq \frac{1}{2} \times \text{length}(x + \mathbb{R}\theta) \cap K$. In other words, we obtain $S_\theta(K)$ by sliding its chords so that their midpoint will be on $\theta^\perp$ and take the union of all resulting chords.

Some basic properties of the Steiner symmetrization are listed below. We leave the details to the reader.

1. Steiner symmetrization preserves convexity: if K is a convex body then $S_\theta(K)$ is also a convex body. In fact, this is the Brunn concavity principle for $k = 1$.
2. $S_\theta(K)$ can be described as follows:
$$S_\theta(K) = \left\{ x + \frac{t_1 - t_2}{2}\theta : x \in P_{\theta^\perp}K, x + t_1\theta \in K, x + t_2\theta \in K \right\}.$$
3. Steiner symmetrization preserves volume: $|S_\theta(K)| = |K|$.
4. Let K_1 and K_2 be two convex bodies. For every $\lambda \in (0, 1)$,
$$S_\theta(\lambda K_1 + (1 - \lambda)K_2) \supseteq \lambda S_\theta(K_1) + (1 - \lambda)S_\theta(K_2).$$

Let F be a k-dimensional subspace of $\mathbb{R}^n$ $(1 \leq k \leq n)$. A well known fact which goes back to Steiner and Schwarz is that for every convex body K one can find a sequence of successive Steiner symmetrizations in directions $\theta \in F$ so that the limiting convex body $\tilde{K}$ has the following property:

> For every $x \in F^\perp$, $\tilde{K} \cap (F + x)$ is a ball with center at x and radius $r(x)$ such that $|\tilde{K} \cap (F + x)| = |K \cap (F + x)|$.

Now, the proof of the theorem is immediate. Convexity of $\tilde{K}$ implies that r is concave on its support, and this shows that f is also concave. □

Proof of the Brunn-Minkowski inequality: Brunn's concavity principle implies the Brunn-Minkowski inequality as follows. If K, T are convex bodies in $\mathbb{R}^n$, we define
$$K_1 = K \times \{0\} \quad \text{and} \quad T_1 = T \times \{1\}$$
in $\mathbb{R}^{n+1}$ and consider their convex hull L. If
$$L(t) = \{x \in \mathbb{R}^n : (x, t) \in L\} \quad , \quad t \in [0, 1]$$
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
$$\left| \frac{K + T}{2} \right|^{1/n} \geq \frac{1}{2}|K|^{1/n} + \frac{1}{2}|T|^{1/n}. \qquad \square$$

Let us give an alternative proof of Theorem 1.2.1, which is due to Gromov and Milman. Their argument proves a more general statement.

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

$$f^{1/\alpha}(\lambda x_1 + \mu x_2) \geq \lambda f^{1/\alpha}(x_1) + \mu f^{1/\alpha}(x_2)$$

for all $x_1, x_2 \in K$ and $\lambda, \mu > 0$ with $\lambda + \mu = 1$.

Lemma 1.2.1 *Let $f, g : K \to \mathbb{R}^+$. If f is α-concave and g is β-concave, then $f \cdot g$ is $(\alpha+\beta)$-concave.*

Proof: Applying Hölder's inequality with $p = (\alpha+\beta)/\alpha$ and $q = (\alpha+\beta)/\beta$, we get

$$\begin{aligned} \lambda\big(f(x_1)g(x_1)\big)^{\frac{1}{\alpha+\beta}} + \mu\big(f(x_2)g(x_2)\big)^{\frac{1}{\alpha+\beta}} &\leq \Big(\lambda f(x_1)^{1/\alpha} + \mu f(x_2)^{1/\alpha}\Big)^{\frac{\alpha}{\alpha+\beta}} \\ &\quad \times \Big(\lambda g(x_1)^{1/\beta} + \mu g(x_2)^{1/\beta}\Big)^{\frac{\beta}{\alpha+\beta}} \\ &\leq \big(f(\lambda x_1 + \mu x_2)\big)^{\frac{1}{\alpha+\beta}} \big(g(\lambda x_1 + \mu x_2)\big)^{\frac{1}{\alpha+\beta}} \end{aligned}$$

for all $x_1, x_2 \in K$ and $\lambda, \mu > 0$ with $\lambda + \mu = 1$, by the α-concavity of f and the β-concavity of g. □

Let now K be a convex body in $\mathbb{R}^n$ and $\theta \in S^{n-1}$. For every $y \in P_{\theta^\perp}(K)$ we write I_y for the set $\{t \in \mathbb{R} : y + t\theta \in K\}$. From the convexity of K it follows that I_y is an interval for every $y \in P_{\theta^\perp}(K)$. Let $f : K \to \mathbb{R}^+$ be a continuous function. We define the **projection** $P_\theta f$ of f with respect to θ by

$$(P_\theta f)(y) := \int_{I_y} f(y + t\theta)dt, \quad y \in P_{\theta^\perp}(K).$$

Lemma 1.2.2 *If f is α-concave, then $P_\theta f$ is $(1+\alpha)$-concave.*

Proof: Concavity of a function is a "two-dimensional" notion, so we may clearly assume that $K \subseteq \mathbb{R}^2$, in which case $P_{\theta^\perp}(K)$ is an interval. Let $y_1, y_2 \in P_{\theta^\perp}(K)$ and write $I_{y_i} = [a_i, b_i]$, $i = 1, 2$. For every $\lambda, \mu > 0$ with $\lambda + \mu = 1$, we set $y_\lambda = \lambda y_1 + \mu y_2$. Then,

$$I_{y_\lambda} \supseteq [\lambda a_1 + \mu a_2, \lambda b_1 + \mu b_2].$$

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 $[y_i + c_i\theta, y_i + b_i\theta]$, $i = 1, 2$ and K'' is the convex hull of the intervals $[y_i + a_i\theta, y_i + c_i\theta]$, $i = 1, 2$, we define $f' = f|_{K'}$ and

$f'' = f|_{K''}$. By the definition of c_i it is not hard to check that if $P_\theta f'$ and $P_\theta f''$ are $(1+\alpha)$-concave then $P_\theta f$ satisfies the $(1+\alpha)$-concavity condition at y_1 and y_2. Thus, we only need to prove that $P_\theta f'$ and $P_\theta f''$ are $(1+\alpha)$-concave.

For every $n \geq 2$ we define partitions $a_i = t_{0,i} < t_{1,i} < \ldots < t_{n-1,i} < t_{n,i} = b_i$ 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 we only need to check that $P_\theta f^k$ is $(1+\alpha)$-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 we see that we have to check the following:

Claim: Let $t_i \in I_{y_i}$ and $d_i > 0$, $i = 1, 2$. Given $\lambda \in (0,1)$, set $y_\lambda = \lambda y_1 + (1-\lambda)y_2$, $t(\lambda) = \lambda t_1 + (1-\lambda)t_2$ and $d(\lambda) = \lambda d_1 + (1-\lambda)d_2$. Then, the function $\lambda \mapsto f(y_\lambda + t(\lambda)\theta) \cdot d(\lambda)$ is $(1+\alpha)$-concave.

This claim follows from Lemma 1.2.1, since $\lambda \mapsto f(y_\lambda + t(\lambda)\theta)$ is α-concave and the linear function $\lambda \mapsto d(\lambda)$ is 1-concave. $\square$

We now finish the proof of Brunn's theorem as follows. The characteristic function of K is constant on K, and hence it is α-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 θ_i. Lemma 1.2.2 shows that the function $x \mapsto |K \cap (F + x)|$ is $(\alpha + k)$-concave on $P_{F^\perp}(K)$, for every $\alpha > 0$. It follows that $f(x) = |K \cap (F + x)|^{1/k}$ is concave. $\square$

1.2.2 Elementary sets

Next, we give a proof of the Brunn-Minkowski inequality for non-empty compact sets in $\mathbb{R}^n$, which 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 1.2.2 *Let A and B be elementary sets in $\mathbb{R}^n$. Then,*

$$|A + B|^{1/n} \geq |A|^{1/n} + |B|^{1/n}.$$

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) to be 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 hyperplane which, without loss of generality, may be described by the equation $x_n=\rho$ for some $\rho\in\mathbb{R}$. We define

$$A^+ = A\cap\{x\in\mathbb{R}^n : x_n\geq\rho\}\quad,\quad A^- = A\cap\{x\in\mathbb{R}^n : x_n\leq\rho\}.$$

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

$$\frac{|A^+|}{|A|}=\frac{|B^+|}{|B|}=:\lambda. \tag{2.1}$$

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,

$$|A+B|\geq|A^++B^+|+|A^-+B^-|.$$

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

$$|A^++B^+|^{1/n}\geq|A^+|^{1/n}+|B^+|^{1/n}\quad\text{and}\quad|A^-+B^-|^{1/n}\geq|A^-|^{1/n}+|B^-|^{1/n}.$$

Taking (2.1) into account we get

$$\begin{aligned}|A+B| &\geq \left(\left(\lambda|A|\right)^{1/n}+\left(\lambda|B|\right)^{1/n}\right)^n+\left(\left((1-\lambda)|A|\right)^{1/n}+\left((1-\lambda)|B|\right)^{1/n}\right)^n\\ &= \lambda\left(|A|^{1/n}+|B|^{1/n}\right)^n+(1-\lambda)\left(|A|^{1/n}+|B|^{1/n}\right)^n\\ &= \left(|A|^{1/n}+|B|^{1/n}\right)^n,\end{aligned}$$

which implies

$$|A+B|^{1/n} \geq |A|^{1/n} + |B|^{1/n}.$$

□

Since every compact set in $\mathbb{R}^n$ can be approximated by a decreasing sequence of elementary sets in the Hausdorff metric, Theorem 1.2.2 is easily extended to non-empty compact sets A and B in $\mathbb{R}^n$.

Theorem 1.2.3 *Let A and B be non-empty compact sets in $\mathbb{R}^n$. Then,*

$$|A+B|^{1/n} \geq |A|^{1/n} + |B|^{1/n}. \qquad \square$$

1.2.3 Induction on the dimension

Our third proof of the Brunn-Minkowski inequality is by induction on the dimension and is due to Kneser and Süss. The argument works only for convex bodies. One of the advantages of the proof is that it clarifies the cases of equality in the convex case. Another one is that suitable modification leads to "stability results".

For the proof we need to introduce the **support function** h_K of a convex body K. For every $u \in S^{n-1}$ we define

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

Observe that the sum $h_K(u) + h_K(-u)$ is the **width** of K in the direction of u. A separation argument shows that two convex bodies K_0 and K_1 coincide if and only if $h_{K_0}(u) = h_{K_1}(u)$ for every $u \in S^{n-1}$.

Theorem 1.2.4 *Let K_0 and K_1 be convex bodies in $\mathbb{R}^n$. For every $\lambda \in [0,1]$,*

$$|(1-\lambda)K_0 + \lambda K_1|^{1/n} \geq (1-\lambda)|K_0|^{1/n} + \lambda |K_1|^{1/n}$$

with equality only if K_0 and K_1 are homothetical.

Proof: We prove the theorem by induction on n. The statement is clearly true when $n = 1$, therefore we may assume that $n \geq 2$ and the theorem has been proved in dimension $n-1$. We may also assume that $|K_0| = |K_1| = 1$ and that the origin is the centroid for both K_0 and K_1. Given $\theta \in S^{n-1}$ we define (for $i = 0, 1$)

$$\begin{aligned} f_i(t) &= |K_i \cap \{y + t\theta : y \in \theta^{\perp}\}|_{n-1} \\ g_i(t) &= |K_i \cap \{y + s\theta : y \in \theta^{\perp},\ s \leq t\}|. \end{aligned}$$

on $\big(-h_{K_i}(-\theta), h_{K_i}(\theta)\big)$. Then,

$$g_i(t) = \int_{-h_{K_i}(-\theta)}^{t} f_i(s)\,ds.$$

Since f_i is continuous and strictly positive, we have

$$g_i'(t) = f_i(t), \qquad -h_{K_i}(-\theta) < t < h_{K_i}(\theta)$$

and g_i is strictly increasing. Now, if $h_i : (0,1) \to \mathbb{R}$ is the inverse function of g_i, we have

$$h_i'(u) = \frac{1}{f_i(h_i(u))}, \qquad 0 < u < 1.$$

Let $K_\lambda = (1-\lambda)K_0 + \lambda K_1$ and write $K_\lambda(u) = K_\lambda \cap \{y + h_\lambda(u)\theta : y \in \theta^\perp\}$ for every $\lambda \in [0,1]$ and $0 < u < 1$, where $h_\lambda = (1-\lambda)h_0 + \lambda h_1$. We can easily check that

$$K_\lambda(u) \supseteq (1-\lambda)K_0(u) + \lambda K_1(u).$$

Therefore, if we make the change of variables $t = h_\lambda(u)$ we get

$$\begin{aligned} |K_\lambda| &= \int_{-h_{K_\lambda}(-\theta)}^{h_{K_\lambda}(\theta)} |K_\lambda \cap (\theta^\perp + t\theta)|_{n-1} dt \\ &\geq \int_0^1 |K_\lambda(u)|_{n-1} h_\lambda'(u) du \\ &\geq \int_0^1 |(1-\lambda)K_0(u) + \lambda K_1(u)|_{n-1} \left(\frac{1-\lambda}{f_0(h_0(u))} + \frac{\lambda}{f_1(h_1(u))} \right) du. \end{aligned}$$

The inductive hypothesis shows that this is greater than or equal to

$$\int_0^1 \left((1-\lambda) f_0(h_0(u))^{\frac{1}{n-1}} + \lambda f_1(h_1(u))^{\frac{1}{n-1}} \right)^{n-1} \left(\frac{1-\lambda}{f_0(h_0(u))} + \frac{\lambda}{f_1(h_1(u))} \right) du.$$

By the arithmetic-geometric means inequality, the integrand is greater than or equal to 1 on $(0,1)$ with equality only if $f_0(h_0(u)) = f_1(h_1(u))$. This shows that $|K_\lambda| \geq 1$ and completes the inductive step.

Suppose that for some $\lambda \in (0,1)$ we have $|(1-\lambda)K_0 + \lambda K_1| = 1$. Then, $f_0(h_0(u)) = f_1(h_1(u))$ for every $u \in (0,1)$, which implies $h_0' = h_1'$. Thus, $h_1 - h_0$ is a constant function. Since 0 is the centroid of K_i,

$$\begin{aligned} 0 &= \int_{K_i} \langle x, \theta \rangle dx = \int t f_i(t)\, dt \\ &= \int_0^1 h_i(u) f_i(h_i(u)) h_i'(u)\, du \\ &= \int_0^1 h_i(u) du, \end{aligned}$$

which shows that $h_0 \equiv h_1$. Then, $h_{K_0}(\theta) = h_{K_1}(\theta)$. Since this should hold for every $\theta \in S^{n-1}$, we must have $K_0 = K_1$. $\square$

1.2.4 Refinements involving hyperplane projections and sections

First, we give a linear refinement of the Brunn-Minkowski inequality for two convex bodies K and T which have projections of the same volume onto some hyperplane.

Lemma 1.2.3 *Let K_0 and K_1 be two convex bodies in $\mathbb{R}^n$ and let $\theta \in S^{n-1}$ be such that $P_{\theta^\perp}(K_0) = P_{\theta^\perp}(K_1) = P$. Then,*

$$|(1-\lambda)K_0 + \lambda K_1| \geq (1-\lambda)|K_0| + \lambda |K_1|$$

for every $\lambda \in (0,1)$.

Proof: Let $K_\lambda = (1-\lambda)K_0 + \lambda K_1$. Note that $P_{\theta^\perp}(K_\lambda) = P$ for all $\lambda \in [0,1]$. We may write

$$K_\lambda = \{y + t\theta : y \in P,\ f_\lambda(y) \leq t \leq g_\lambda(y)\},$$

where f_λ is convex on P, g_λ is concave on P, and $f_\lambda \leq g_\lambda$. We easily check that $g_\lambda \geq (1-\lambda)g_0 + \lambda g_1$ and $f_\lambda \leq (1-\lambda)f_0 + \lambda f_1$ on P, hence

$$\begin{aligned}
|K_\lambda| &= \int_P \big(g_\lambda(y) - f_\lambda(y)\big)dy \\
&\geq \int_P \big((1-\lambda)g_0(y) + \lambda g_1(y) - (1-\lambda)f_0(y) - \lambda f_1(y)\big)dy \\
&= (1-\lambda)\int_P \big(g_0(y) - f_0(y)\big)dy + \lambda \int_P \big(g_1(y) - f_1(y)\big)dy \\
&= (1-\lambda)|K_0| + \lambda |K_1|. \qquad \square
\end{aligned}$$

Lemma 1.2.3 and a symmetrization argument give a stronger result.

Theorem 1.2.5 *Let K_0 and K_1 be two convex bodies in $\mathbb{R}^n$ and let $\theta \in S^{n-1}$ be such that $|P_\theta(K_0)| = |P_\theta(K_1)|$. Then,*

$$|(1-\lambda)K_0 + \lambda K_1| \geq (1-\lambda)|K_0| + \lambda |K_1|$$

for every $\lambda \in (0,1)$.

Proof: We first symmetrize K_0 and K_1 with respect to θ. If $K_0' = S_\theta(K_0)$ and $K_1' = S_\theta(K_1)$, then $|K_0| = |K_0'|$, $|K_1| = |K_1'|$ and

$$(1-\lambda)K_0' + \lambda K_1' \subseteq S_\theta\big((1-\lambda)K_0 + \lambda K_1\big).$$

Since Steiner symmetrization preserves volume, it suffices to prove that

$$|(1-\lambda)K_0' + \lambda K_1'| \geq (1-\lambda)|K_0\prime| + \lambda |K_1'|. \tag{2.2}$$

By a sequence of Steiner symmetrizations in directions $\xi \in \theta^{\perp}$, we obtain two rotation bodies K_0'' and K_1'' which have the same projection onto $\theta^{\perp}$: a centered ball of volume $V = |P_{\theta^{\perp}}(K_0)| = |P_{\theta^{\perp}}(K_1)|$. Lemma 1.2.3 shows that

$$|(1-\lambda)K_0'' + \lambda K_1''| \geq (1-\lambda)|K_0''| + \lambda |K_1''|,$$

and (2.2) follows because

$$|(1-\lambda)K_0' + \lambda K_1'| \geq |(1-\lambda)K_0'' + \lambda K_1''|$$

by the same reasoning as in the beginning of the proof. □

For the second result we need to introduce some notation: Let f and g be two bounded non-negative Borel measurable functions on $\mathbb{R}$ which vanish outside a finite interval. We define

$$h(r,s) = f(r) + g(s) \quad \text{if} \quad f(r)g(s) > 0$$

and $h(r,s) = 0$ otherwise. Then, we set

$$m(t) = \sup\{h(r,s) \mid r + s = t\}.$$

We also write $\gamma = \sup_r f(r)$ and $\delta = \sup_s g(s)$.

Lemma 1.2.4 *With this notation,*

$$\int_{\mathbb{R}} m^p(t)dt \geq (\gamma+\delta)^p \left[\frac{1}{\gamma^p}\int_{\mathbb{R}} f^p(t)dt + \frac{1}{\delta^p}\int_{\mathbb{R}} g^p(t)dt\right]$$

for every $p > 0$.

Proof: If $u : \mathbb{R} \to \mathbb{R}^+$ is a measurable function, we set

$$A_u(\alpha) = \{s : u(s) \geq \alpha\}.$$

By the definition of m we easily check that

$$A_m((\gamma+\delta)t) \supseteq A_f(\gamma t) + A_g(\delta t)$$

for every $t \in (0,1)$. One dimensional Brunn-Minkowski inequality implies that

$$|A_m((\gamma+\delta)t)| \geq |A_f(\gamma t)| + |A_g(\delta t)|.$$

It follows that

$$\begin{aligned}
\int_{\mathbb{R}} m^p &= \int_0^{\gamma+\delta} |A_m(\alpha)| p\alpha^{p-1} d\alpha \\
&= (\gamma+\delta)^p \int_0^1 |A_m((\gamma+\delta)t)| pt^{p-1} dt \\
&\geq (\gamma+\delta)^p \left(\int_0^1 |A_f(\gamma t)| pt^{p-1} dt + \int_0^1 |A_g(\delta t)| pt^{p-1} dt\right) \\
&= (\gamma+\delta)^p \left[\frac{1}{\gamma^p}\int_{\mathbb{R}} f^p(t)dt + \frac{1}{\delta^p}\int_{\mathbb{R}} g^p(t)dt\right]. \quad \square
\end{aligned}$$

Theorem 1.2.6 *Let K and T be two convex bodies in $\mathbb{R}^n$. Fix $\theta \in S^{n-1}$ and set*

$$\gamma = \sup_r |K \cap (\theta^\perp + r\theta)|^{\frac{1}{n-1}} \quad , \quad \delta = \sup_s |T \cap (\theta^\perp + s\theta)|^{\frac{1}{n-1}}.$$

Then,

$$|K + T| \geq (\gamma + \delta)^{n-1} \left(\frac{|K|}{\gamma^{n-1}} + \frac{|T|}{\delta^{n-1}} \right).$$

Proof: Consider the functions

$$f(r) = |K \cap (\theta^\perp + r\theta)|^{\frac{1}{n-1}} \quad , \quad g(s) = |T \cap (\theta^\perp + s\theta)|^{\frac{1}{n-1}}$$

and

$$u(t) = |(K + T) \cap (\theta^\perp + t\theta)|^{\frac{1}{n-1}}.$$

An application of the Brunn-Minkowski inequality in dimension $n-1$ shows that $u \geq m$, where m is defined as in Lemma 1.2.4. Observe that

$$\int_{\mathbb{R}} f^{n-1} = |K| \quad , \quad \int_{\mathbb{R}} g^{n-1} = |T| \quad \text{and} \quad \int_{\mathbb{R}} u^{n-1} = |K + T|.$$

Applying Lemma 1.2.4 with $p = n - 1$ we conclude the proof. $\square$

This inequality is formally stronger than the Brunn-Minkowski inequality. Also, if the bodies K and T have equal maximal hyperplane sections in some direction, we obtain a linear refinement of the Brunn-Minkowski inequality.

Corollary 1.2.1 *Let K and T be two convex bodies in $\mathbb{R}^n$. If*

$$\sup_r |K \cap (\theta^\perp + r\theta)| = \sup_s |T \cap (\theta^\perp + s\theta)|$$

for some $\theta \in S^{n-1}$, then

$$\left| \frac{K + T}{2} \right| \geq \frac{|K| + |T|}{2}. \qquad \square$$

1.3 Volume preserving transformations

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 costant determinant equal to $|K|/|T|$. In this section we describe two such maps, the Knöthe map and the Brenier map. Applying each one of them we may obtain alternative proofs of the Brunn-Minkowski inequality.

1.3.1 Knöthe map

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

Theorem 1.3.1 *There exists a map $\phi : K \to T$ with the following properties:*

(a) ϕ is triangular: the i-th coordinate function of ϕ 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 of ϕ is constant. More precisely, for every $x \in K$

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

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

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

That is, 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 ϕ_1 is defined on K but $\phi_1(x)$ depends only on the first coordinate of $x \in K$. Also,

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

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

$$\frac{1}{|K_{(x_1,\ldots,x_{j-1})}|} \int_{-\infty}^{x_j} |K_{(x_1,\ldots,x_{j-1},s_j)}|_{n-j} ds_j$$

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

It is clear that

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

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

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

By construction, ϕ has properties (a) and (b). □

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

Using the Knöthe 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 Knöthe map $\phi : K \to T$. It is clear that

$$(I + \phi)(K) \subseteq K + \phi(K) = K + T,$$

and hence,

$$\begin{aligned}
|K + T| &\geq \int_{(I+\phi)(K)} dx = \int_K |J(I + \phi)(x)| dx \\
&= \int_K \prod_{j=1}^n \left(1 + \frac{\partial \phi_j}{\partial x_j}(x)\right) dx \\
&\geq \int_K \left(1 + \left(\prod_{j=1}^n \frac{\partial \phi_j}{\partial x_j}(x)\right)^{1/n}\right)^n dx \\
&= |K| \left(1 + \left(\frac{|T|}{|K|}\right)^{1/n}\right)^n \\
&= \left(|K|^{1/n} + |T|^{1/n}\right)^n. \quad \square
\end{aligned}$$

As an example of application, we give a direct proof of Theorem 1.2.5, which avoids the use of symmetrization.

Theorem 1.3.2 *Let K and T be two convex bodies in $\mathbb{R}^n$. Assume that $|P_F(K)| = |P_F(T)|$ for some $(n-1)$-dimensional subspace F of $\mathbb{R}^n$. Then,*

$$|\lambda K + (1 - \lambda)T| \geq \lambda |K| + (1 - \lambda)|T|$$

for every $\lambda \in (0, 1)$.

Proof: We may assume that $F = \langle e_1, \ldots, e_{n-1} \rangle$, where $\{e_1, \ldots, e_n\}$ is the standard orthonormal basis in $\mathbb{R}^n$. We may also replace K and T by their interiors. We consider the Knöthe map $\phi : P_F(K) \to P_F(T)$. Since $|P_F(K)| = |P_F(T)|$, we have

$$J(\phi)(x) = 1, \quad x \in P_F(K).$$

For every $x \in P_F(K)$, we define

$$K(x) = \{y \in \mathbb{R} : (x, y) \in K\}.$$

$T(w)$ is defined in the same way for every $w \in P_F(T)$. Given $x \in P_F(K)$, we consider the "ratio of lengths" preserving linear map $\psi_x : K(x) \to T(\phi(x))$. Then,

$$\psi_x'(y) = \frac{|T(\phi(x))|}{|K(x)|}$$

for all $x \in P_F(K)$ and $y \in K(x)$. Finally, we define $f : K \to T$ as follows: every $z \in K$ can be written in the form $z = (x, y)$, where $x = P_F(z) \in P_F(K)$ and $y = z - P_F(z) \in K(x)$. We then set

$$f(x, y) = (\phi(x), \psi_x(y)).$$

It is not hard to check that f is one-to-one from K onto T, and hence,

$$\lambda K + (1 - \lambda)T \supseteq (\lambda I + (1 - \lambda)f)(K).$$

It follows that

$$\begin{aligned} |\lambda K + (1 - \lambda)T| &\geq |(\lambda I + (1 - \lambda)f)(K)| = \int_{(\lambda I + (1-\lambda)f)(K)} dw \\ &= \int_K J(\lambda I + (1 - \lambda)f)(z)dz. \end{aligned}$$

By the definition of f, we see that

$$J(\lambda I + (1 - \lambda)f)(z) = J(\lambda I + (1 - \lambda)\phi)(x) \cdot (\lambda + (1 - \lambda)\psi_x'(y)$$

for all $z = (x, y) \in K$, which implies

$$|\lambda K + (1 - \lambda)T| \geq \int_{P_F(K)} J(\lambda I + (1 - \lambda)\phi)(x) \left(\lambda|K(x)| + (1 - \lambda)|T(\phi(x))|\right) dx.$$

By the arithmetic-geometric means inequality we see that $J(\lambda I + (1 - \lambda)\phi)(x) \geq 1$ on $P_F(K)$. It follows that

$$\begin{aligned} |\lambda K + (1 - \lambda)T| &\geq \int_{P_F(K)} \left(\lambda|K(x)| + (1 - \lambda)|T(\phi(x))|\right) dx \\ &= \lambda|K| + (1 - \lambda) \int_{\phi^{-1}(P_F(T))} |T(\phi(x))|dx \\ &= \lambda|K| + (1 - \lambda) \int_{P_F(T)} |T(u)|du \\ &= \lambda|K| + (1 - \lambda)|T|. \end{aligned}$$

This completes the proof. $\square$

1.3.2 Brenier map

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 μ, $\nu \in \mathcal{P}(\mathbb{R}^n)$. We say that a probability measure $\gamma \in \mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n)$ has marginals μ and ν if for every bounded Borel measurable $f, g : \mathbb{R}^n \to \mathbb{R}$ 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

$$\int_{\mathbb{R}^n} g(y)d\nu(y) = \int_{\mathbb{R}^n\times\mathbb{R}^n} g(y)d\gamma(x,y).$$

Definition Let $G \subset \mathbb{R}^n \times \mathbb{R}^n$. 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

$$\langle y_1, x_2 - x_1\rangle + \langle y_2, x_3 - x_2\rangle + \ldots + \langle y_m, x_1 - x_m\rangle \leq 0.$$

Proposition 1.3.1 *Let μ and ν be Borel probability measures on $\mathbb{R}^n$. There exists a joint probability measure γ on $\mathbb{R}^n \times \mathbb{R}^n$ which has cyclically monotone support and marginals μ, ν.*

The proof of the Proposition follows by the next discrete lemma: then, one has to show that if γ_n has cyclically monotone support and marginals μ_n, ν_n which converge in the weak-$*$ topology to μ and ν, there exists a weak-$*$ subsequential limit γ which has cyclically monotone support and μ, ν as its marginals. $\square$

Lemma 1.3.1 *Let $x_i, y_i \in \mathbb{R}^n$, $i = 1, \ldots, m$ and consider the measures*

$$\mu = \frac{1}{m}\sum_{i=1}^m \delta_{x_i} \quad \textit{and} \quad \nu = \frac{1}{m}\sum_{i=1}^m \delta_{y_i}.$$

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

Proof: For every permutation σ 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 γ_σ has marginals μ and ν for every σ. Let

$$F(\sigma) = \sum_{i=1}^m \langle x_{\sigma(i)}, y_i\rangle.$$

We will show that if $F(\sigma)$ is maximal, then the support $(x_{\sigma(i)}, y_i)$ of γ_σ 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 $(x_{i_s}, y_{i_s}) \in G$. If σ 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

$$\begin{aligned} 0 \geq F(\sigma) - F(I) &= \sum_{s=1}^{k} \left(\langle x_{\sigma(i_s)}, y_{i_s} \rangle - \langle x_{i_s}, y_{i_s} \rangle \right) \\ &= \langle x_{i_2} - x_{i_1}, y_{i_1} \rangle + \langle x_{i_3} - x_{i_2}, y_{i_2} \rangle + \ldots + \langle x_{i_1} - x_{i_k}, y_{i_k} \rangle. \end{aligned}$$

This proves the lemma. □

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a **proper** convex function. This means that f takes values in $\mathbb{R} \cup \{+\infty\}$, it is not identically $+\infty$ and its restriction onto any line in $\mathbb{R}^n$ is a convex function. The **domain** of f is the convex set $\mathrm{dom}(f) = \{x : f(x) < +\infty\}$. If A is the interior of $\mathrm{dom}(f)$, then f is continuous on A and ∇f exists almost everywhere in A. Moreover, we can modify the values of f on the boundary of A so that f will become lower semicontinuous (then, we say that f is **closed**).

The **subdifferential** of f is the set

$$\partial(f) = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : f(z) \geq f(x) + \langle y, z - x \rangle, \ z \in \mathbb{R}^n\}.$$

In other words, the subdifferential parametrizes the supporting hyperplanes of f. For each $x \in A$, the set $\partial(f)(x) = \{y : (x, y) \in \partial(f)\}$ is a closed and bounded convex set. It is easy to see that $\nabla f(x)$ exists if and only if $\partial(f)(x) = \{\nabla f(x)\}$.

Proposition 1.3.2 *Let $G \subset \mathbb{R}^n \times \mathbb{R}^n$. Then, G is contained in the subdifferential of a proper convex function $f : \mathbb{R}^n \to \mathbb{R}$ 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 $(x_i, y_i) \in \partial(f)$, $i = 1, \ldots, m$. Then,

$$\begin{aligned} \langle y_1, x_2 - x_1 \rangle &\leq f(x_2) - f(x_1) \\ \langle y_2, x_3 - x_2 \rangle &\leq f(x_3) - f(x_2) \\ & \\ \langle y_m, x_1 - x_m \rangle &\leq f(x_1) - f(x_m) \end{aligned}$$

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

$$\langle y_1, x_2 - x_1 \rangle + \langle y_2, x_3 - x_2 \rangle + \ldots + \langle y_m, x_1 - x_m \rangle \leq 0.$$

It follows that every $G \subseteq \partial(f)$ is cyclically monotone.

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

$$f(x) = \sup\left\{ \langle y_m, x - x_m \rangle + \langle y_{m-1}, x_m - x_{m-1} \rangle + \ldots + \langle y_0, x_1 - x_0 \rangle \right\},$$

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

$$t + \langle z - x, y \rangle < f(z)$$

for every $t < f(x)$ and every $z \in \mathbb{R}^n$. This implies that $(x, y) \in \partial(f)$. Since $t < f(x)$, 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 f again,

$$\begin{aligned} f(z) &\geq \langle y, z - x \rangle + \langle y_m, x - x_m \rangle + \cdots + \langle y_0, x_1 - x_0 \rangle \\ &> \langle y, z - x \rangle + t. \end{aligned}$$

This completes the proof. $\square$

Let μ, $\nu \in \mathcal{P}(\mathbb{R}^n)$. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a measurable function which is defined μ-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** μ to ν 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

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

Theorem 1.3.3 *Let μ, $\nu \in \mathcal{P}(\mathbb{R}^n)$ and assume that μ is absolutely continuous with respect to Lebesgue measure. Then, there exists a convex function $f : \mathbb{R}^n \to \mathbb{R}$ such that $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ is defined μ-almost everywhere, and $(\nabla f)\mu = \nu$.*

Proof: Proposition 1.3.1 shows that there exists a probability measure γ on $\mathbb{R}^n \times \mathbb{R}^n$ which has cyclically monotone support and marginals μ, ν. By Proposition 1.3.2, the support of γ is contained in the subdifferential of a proper convex function $f : \mathbb{R}^n \to \mathbb{R}$.

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

$$\begin{aligned} \int g(y)d\nu(y) &= \int g(y)d\gamma(x,y) \\ &= \int g(\nabla f(x))d\gamma(x,y) \\ &= \int g(\nabla f(x))d\mu(x), \end{aligned}$$

which shows that $(\nabla f)\mu = \nu$. □

A remarkable property of the Brenier map ∇f is that it is uniquely determined. Assume that μ is the normalized Lebesgue measure on some convex body K and ν is the normalized Lebesgue measure on some other convex body T. Regularity results of Caffarelli show that in this case f may be assumed twice continuously differentiable.

Theorem 1.3.4 *Let K and T be open convex bodies in $\mathbb{R}^n$. There is a convex function $f \in C^2(K)$ such that $\psi = \nabla f : K \to T$ is one to one, onto and volume preserving.* □

We can now show the Brunn-Minkowski inequality using ψ. It is clear that the Jacobian $J(\psi) = \text{Hess}f$ is a symmetric positive definite matrix for every $x \in K$. Since $(I + \psi)(K) \subseteq K + T$,

$$|K+T| \geq \int_K |J(I+\psi)(x)|dx = \int_K |\left(I + \text{Hess}f\right)(x)|dx = \int_K \prod_{i=1}^n (1+\lambda_i(x))dx,$$

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

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

This proof has the advantage of providing a description for the equality cases: either K or T must be a point, or K must be homothetical to T. □

1.4 Functional forms

A **functional form** of the Brunn-Minkowski inequality is an integral inequality which reduces to the Brunn-Minkowski inequality by appropriate choice of the functions involved. Such functional inequalities can be applied to different contexts: as an example, in this Section we will see how the Prékopa-Leindeler inequality may be applied to yield the logarithmic Sobolev inequality and several important concentration results in Gauss space.

1.4.1 Prékopa-Leindler inequality

The inequality of Prékopa and Leindler is the following statement.

Theorem 1.4.1 *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 for every $x, y \in \mathbb{R}^n$*

$$h(\lambda x + (1-\lambda)y) \geq f(x)^\lambda g(y)^{1-\lambda}. \tag{4.1}$$

Then,

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

Proof: The proof goes by induction on the dimension n.

(a) $n = 1$: We may assume that f and g are continuous and strictly positive. We define $x, y : (0,1) \to \mathbb{R}$ by the equations

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

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

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

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

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

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

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

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

$$\begin{aligned}
\int h &= \int_0^1 h(z(t)) z'(t) dt \\
&\geq \int_0^1 h(\lambda x(t) + (1-\lambda) y(t)) (x'(t))^\lambda (y'(t))^{1-\lambda} dt \\
&\geq \int_0^1 f^\lambda(x(t)) g^{1-\lambda}(y(t)) \left(\frac{\int f}{f(x(t))}\right)^\lambda \left(\frac{\int g}{g(y(t))}\right)^{1-\lambda} dt \\
&= \left(\int f\right)^\lambda \left(\int g\right)^{1-\lambda}.
\end{aligned}$$

(b) *Inductive step:* We assume that $n \geq 2$ and 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 (4.1) it follows that if $x, y \in \mathbb{R}^{n-1}$ and $s_0, s_1 \in \mathbb{R}$ then

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

and our inductive hypothesis gives

$$\begin{aligned} H(\lambda s_1 + (1-\lambda)s_0) &:= \int_{\mathbb{R}^{n-1}} h_{\lambda s_1 + (1-\lambda)s_0} \\ &\geq \left(\int_{\mathbb{R}^{n-1}} f_{s_1}\right)^\lambda \left(\int_{\mathbb{R}^{n-1}} g_{s_0}\right)^{1-\lambda} =: F^\lambda(s_1) G^{1-\lambda}(s_0). \end{aligned}$$

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

$$\int h = \int_{\mathbb{R}} H \geq \left(\int_{\mathbb{R}} F\right)^\lambda \left(\int_{\mathbb{R}} G\right)^{1-\lambda} = \left(\int f\right)^\lambda \left(\int g\right)^{1-\lambda}. \qquad \square$$

Proof of the Brunn-Minkowski inequality: Let K and T be non-empty compact subsets of $\mathbb{R}^n$, and $\lambda \in (0,1)$. We define $f = \chi_K$, $g = \chi_T$, and $h = \chi_{\lambda K + (1-\lambda)T}$. It is easily checked that the assumptions of Theorem 1.4.1 are satisfied, therefore

$$|\lambda K + (1-\lambda)T| = \int h \geq \left(\int f\right)^\lambda \left(\int g\right)^{1-\lambda} = |K|^\lambda |T|^{1-\lambda}. \qquad \square$$

In order to state our next result, we introduce some notation: If $p > 0$ and $\lambda \in (0,1)$, for all $x, y > 0$ we set

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

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

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

By Hölder's inequality, if $x, y, z, w \geq 0$, $a, b, \gamma > 0$ and $\frac{1}{a} + \frac{1}{b} = \frac{1}{\gamma}$, then

$$M_a^\lambda(x, y) \cdot M_b^\lambda(z, w) \geq M_\gamma^\lambda(xz, yw). \tag{4.2}$$

Proposition 1.4.1 *Let $f, g, h : \mathbb{R}^n \to \mathbb{R}^+$ be measurable functions, and let $p > 0$, $\lambda \in (0,1)$. We assume that f and g are integrable, and for every $x, y \in \mathbb{R}^n$*

$$h(\lambda x + (1-\lambda)y) \geq M_p^\lambda(f(x), g(y)). \tag{4.3}$$

Then,

$$\int_{\mathbb{R}^n} h \geq M^\lambda_{\frac{p}{pn+1}}\left(\int_{\mathbb{R}^n} f, \int_{\mathbb{R}^n} g\right).$$

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

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

Then, z is strictly increasing, and

$$z'(t) = \lambda x'(t) + (1 - \lambda)y'(t) = M_1^{\lambda}(x'(t), y'(t)).$$

Hence, using (4.2) and (4.3) we can estimate the integral of h making the change of variables $s = z(t)$, as follows:

$$\begin{aligned}
\int h &= \int_0^1 h(z(t))z'(t)dt \\
&\geq \int_0^1 h(\lambda x(t) + (1 - \lambda)y(t))M_1^{\lambda}(x'(t), y'(t))dt \\
&\geq \int_0^1 M_p^{\lambda}(f(x(t)), g(y(t)))M_1^{\lambda}(x'(t), y'(t))dt \\
&\geq \int_0^1 M_{\frac{p}{p+1}}^{\lambda}(f(x(t))x'(t), g(y(t))y'(t))\, dt \\
&= M_{\frac{p}{p+1}}^{\lambda}\left(\int f, \int g\right).
\end{aligned}$$

The inductive step is exactly as in Theorem 1.4.1. □

Remark: Using Proposition 1.4.1 we may give an alternative proof of Lemma 1.2.2. The claim was the following: K is a two-dimensional convex body and $f : K \to \mathbb{R}^+$ is an α-concave function. If

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

then Pf is $(1 + \alpha)$-concave.

Proof: Let $F_y(t) = \chi_K(y, t)f(y, t)$, $y \in PK$. Then, for all $y, z \in PK$ we have

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

by the α-concavity of f and the convexity of K. The claim follows from Proposition 1.4.1. □

1.4.2 The logarithmic Sobolev inequality

Let γ_n denote the standard Gaussian measure in $\mathbb{R}^n$ with density

$$\gamma_n(x) = (2\pi)^{-n/2}\exp(-|x|^2).$$

Maurey observed that the Prékopa-Leindler inequality may be used to give a simple proof of the approximate isoperimetric inequality in the Gauss space $(\mathbb{R}^n, \gamma_n)$.

Theorem 1.4.2 *Let A be a non-empty Borel subset of $\mathbb{R}^n$. Then,*

$$\int_{\mathbb{R}^n} e^{d(x,A)^2/4} d\gamma_n(x) \leq \frac{1}{\gamma_n(A)},$$

where $d(x,A) = \inf\{|x-y| : y \in A\}$. Therefore, if $\gamma_n(A) = \frac{1}{2}$ we have

$$1 - \gamma_n(A_t) \leq 2\exp(-t^2/4)$$

for every $t > 0$, where $A_t = \{y \mid d(y,A) \leq t\}$.

Proof: Consider the functions

$$f(x) = e^{d(x,A)^2/4}\gamma_n(x) \quad , \quad g(x) = \chi_A(x)\gamma_n(x) \quad , \quad m(x) = \gamma_n(x).$$

For every $x \in \mathbb{R}^n$ and $y \in A$ we see that

$$\begin{aligned} (2\pi)^n f(x)g(y) &= e^{d(x,A)^2/4}e^{-|x|^2/2}e^{-|y|^2/2} \leq \exp\big(\frac{|x-y|^2}{4} - \frac{|x|^2}{2} - \frac{|y|^2}{2}\big) \\ &= \exp\big(-\frac{|x+y|^2}{4}\big) = \Big(\exp(-\frac{1}{2}|\frac{x+y}{2}|^2)\Big)^2 \\ &= (2\pi)^n\big(m(\frac{x+y}{2})\big)^2, \end{aligned}$$

by the parallelogram law and the fact that $d(x,A) \leq |x-y|$. Since $g(y) = 0$ whenever $y \notin A$, this implies that f, g, m satisfy the assumptions of the Prékopa-Leindler inequality with $\lambda = 1/2$. Therefore,

$$\Big(\int e^{d(x,A)^2/4}\gamma_n(dx)\Big)\gamma_n(A) = \Big(\int f\Big)\Big(\int g\Big) \leq \Big(\int m\Big)^2 = 1.$$

This proves the first assertion of the theorem. For the second one, observe that if $\gamma_n(A) = \frac{1}{2}$ then

$$e^{t^2/4}\gamma_n(x : d(x,A) \geq t) \leq \int e^{d(x,A)^2/4}\gamma_n(dx) \leq \frac{1}{\gamma_n(A)} = 2.$$

This shows that $\gamma_n(A_t^c) \leq 2\exp(-t^2/4)$. $\square$

Bobkov and Ledoux exploited this idea to deduce the logarithmic Sobolev inequality in a more general setting: Let $E = (\mathbb{R}^n, \|\cdot\|)$ be a finite dimensional normed space and let $E^* = (\mathbb{R}^n, \|\cdot\|_*)$ be its dual. Consider a probability measure μ on E with density $\rho e^{-V(x)}$, where V is a convex function defined on some open convex subset Ω of E. Assume further that there is a constant $c > 0$ with the following property: for any $s, t > 0$ with $s + t = 1$ and any $x, y \in \Omega$,

$$tV(x) + sV(y) - V(tx + sy) \geq \frac{cts}{2}\|x - y\|^2.$$

Let $f \in C^\infty(\Omega)$. The **entropy** of f^2 with respect to μ is defined by

$$\mathrm{Ent}_\mu(f^2) = \int f^2 \log f^2 d\mu - \int f^2 d\mu \cdot \log \int f^2 d\mu.$$

The **logarithmic Sobolev inequality** is the following statement.

Theorem 1.4.3 *For all $f \in C^\infty(\Omega)$,*

$$\mathrm{Ent}_\mu(f^2) \leq \frac{2}{c} \int \|\nabla f\|_*^2 d\mu.$$

Proof: We may set $f^2 = e^g$ where $g \in C_b^\infty(\Omega)$, that is, g has compact support in Ω and bounded partial derivatives. Let $t, s > 0$ with $t + s = 1$. Consider the functions $u(x) = e^{g(x)/t - V(x)}$, $v(y) = e^{-V(y)}$ and $w(z) = e^{g_t(z) - V(z)}$, where

$$g_t(z) = \sup\{g(x) - [tV(x) + sV(y) - V(tx + sy)] : z = tx + sy, x, y \in \Omega\}.$$

Then, $u, v, w : \mathbb{R}^n \to \mathbb{R}^+$ are measurable and if $z = tx + sy$ we have $w(z) \geq u(t)^t u(y)^s$. Indeed,

$$\begin{aligned}
u^t(x)v^s(y) &= \exp(g(x) - tV(x))\exp(-sV(y)) \\
&= \exp(g(x) - tV(x) - sV(y)) \\
&= \exp(g(x) - [tV(x) + sV(y) - V(tx + sy)] - V(tx + sy)) \\
&\leq \exp\big(\sup_{z = tx + sy}\{g(x) - [tV(x) + sV(y) - V(tx + sy)]\} - V(tx + sy)\big) \\
&= w(z).
\end{aligned}$$

An application of the Prékopa-Leindler inequality gives

$$\begin{aligned}
\int e^{g_t} d\mu &= \rho \int e^{g_t(z) - V(z)} dx \\
&\geq \left(\rho \int e^{g(x)/t - V(x)} dx\right)^t \left(\rho \int e^{-V(x)}\right)^s \\
&= \left(\int e^{g/t} d\mu\right)^t.
\end{aligned}$$

If we develop the right hand side around $t = 1$ we get

$$\left(\int e^{g/t} d\mu\right)^t = \int e^g d\mu + s\mathrm{Ent}_\mu(e^g) + O(s^2).$$

To see this, let $h(t) = (\int e^{g/t} d\mu)^t = e^{t \log \int e^{g/t} d\mu}$. Then,

$$h(t) = h(1) + h'(1)(t-1) + O((t-1)^2) = h(1) - h'(1)s + O(s^2).$$

On the other hand,

$$h'(t) = \left(\int e^{g/t} d\mu\right)^t \left(\log \int e^{g/t} d\mu - \frac{\int e^g g d\mu}{t \int e^g d\mu}\right),$$

and hence $h'(1) = -\mathrm{Ent}_\mu(e^g)$.

We now pass to the left hand side. By assumption,

$$g(x) - [tV(x) + sV(y) - V(tx+sy)] \leq g(x) - \frac{cts}{2}\|x-y\|^2$$

for all $x, y \in \Omega$. The definition of g_t shows that

$$g_t(z) \leq \sup\left\{g(x) - \frac{cts\|x-y\|^2}{2} : z = tx + sy,\, x, y \in \Omega\right\}.$$

Now, $z = tx + sy$ can be written in the form $x = z + \frac{s(z-y)}{t}$. If we set $h = z - y$ and $\eta = s/t$, then

$$g_t(z) \leq \sup_{h \in E} \left\{g(z + \eta h) - \frac{c\eta\|h\|^2}{2}\right\}.$$

Observe that

$$g(z + \eta h) = g(z) + \eta\langle \nabla g(z), h\rangle + \eta^2 O(\|h\|^2),$$

independently from z. Hence, we can write

$$\begin{aligned} g_t(z) &\leq \sup\left\{g(z) + \eta\langle \nabla g(z), h\rangle + \eta^2 O(\|h\|^2) - \frac{c\eta}{2}\|h\|^2\right\} \\ &= g(z) + \eta \sup\left\{\langle \nabla g(z), h > -(\frac{c}{2} - \eta C)\|h\|^2\right\}. \end{aligned}$$

Every $h \in \mathbb{R}^n$ can be written in the form $h = \lambda e$ where $\lambda \geq 0$ and $\|e\| = 1$. If we set $\theta = c - 2\eta C$, we have

$$\begin{aligned} g_t(z) &\leq g(z) + \eta \sup_{\lambda \geq 0} \sup_{\|e\|=1} \left\{\lambda\langle \nabla g(z), e\rangle - (\frac{c}{2} - \eta C)\lambda^2\right\} \\ &= g_t(z) \leq g(z) + \eta \sup\left\{\lambda\|\nabla g(z)\|_* - \theta\frac{\lambda^2}{2} : \lambda \geq 0\right\} \\ &= g(z) + \eta\frac{\|\nabla g(z)\|_*^2}{2\theta}. \end{aligned}$$

Now,
$$\frac{\eta}{2\theta} = \frac{\eta}{2c} + O(\eta^2), \qquad |\eta| < \frac{c}{2C}$$
and since $\|\nabla g(z)\|_*$ is bounded, we get
$$\frac{\eta}{2\theta}\|\nabla g(z)\|_*^2 = \frac{\eta}{2c}\|\nabla g(z)\|_*^2 + O(\eta^2).$$
Therefore,
$$g_t(z) = g(z) + \frac{\eta}{2c}\|\nabla g(z)\|_*^2 + O(\eta^2).$$
Applying Taylor's theorem to $x \mapsto e^x$ at x we have
$$e^y = e^x + e^x(y-x) + O((y-x)^2).$$
This shows that
$$\begin{aligned} e^{g_t(z)} &= e^{g(z)+\frac{\eta}{2c}\|\nabla g(z)\|_*^2+O(\eta^2)} \\ &= e^{g(z)} + e^{g(z)}\frac{\eta}{2c}\|\nabla g(z)\|_*^2 + O(\eta^2), \end{aligned}$$
and hence,
$$\begin{aligned} \int e^{g_t(z)}d\mu &= \int e^{g(z)}d\mu + \frac{\eta}{2c}\int \|\nabla g(z)\|_*^2 e^{g(z)}d\mu + O(\eta^2) \\ &= \int e^{g(z)}d\mu + \frac{\eta}{2c}\int \|\nabla g(z)\|_*^2 e^{g(z)}d\mu + O(\eta^2). \end{aligned}$$
Going back to the inequality $\left(\int e^{g/t}d\mu\right)^t \leq \int e^{g_t}d\mu$ we get
$$s\mathrm{Ent}_\mu(e^g) = \frac{\eta}{2c}\int \|\nabla g\|_*^2 e^g d\mu + O(s^2),$$
or
$$\mathrm{Ent}_\mu(e^g) = \frac{1}{2c(1-s)}\int \|\nabla g\|_*^2 e^g d\mu + O(s).$$
Letting $s \to 0$ we arrive at
$$\mathrm{Ent}_\mu(e^g) \leq \frac{1}{2c}\int \|\nabla g\|_*^2 e^g d\mu. \tag{4.4}$$
From $f = e^{g/2}$ we see that $2\nabla f = e^{g/2}\nabla g$, which implies $4\|\nabla f\|^2 = e^g\|g\|^2$. Then, (4.4) gives
$$\mathrm{Ent}_\mu(f^2) = \mathrm{Ent}_\mu(e^g) \leq \frac{1}{2c}\int 4\|f\|^2 d\mu = \frac{2}{c}\int \|\nabla f\|^2 d\mu. \qquad \square$$

Remark: If $\mu = \gamma_n$, then $V(x) = |x|^2/2$ and the condition in Theorem 1.4.3 is satisfied as an equality with $c = 2$. Since $\|\cdot\|_* = |\cdot|$, the logarithmic Sobolev inequality takes the form
$$\mathrm{Ent}(f^2) \leq \int |\nabla f|^2 d\gamma_n$$
in Gauss space.

1.4.3 Concentration of measure in Gauss space

In this subsection we give two applications of the logarithmic Sobolev inequality. The first one is an inequality of Poincaré.

Theorem 1.4.4 *Let γ_n be the standard Gaussian measure on $\mathbb{R}^n$. If f and $|\nabla f|$ are integrable, then*

$$\int_{\mathbb{R}^n} f^2 d\gamma_n - \left(\int_{\mathbb{R}^n} f d\gamma_n\right)^2 \leq \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n.$$

Proof: By homogenuity we may assume that

$$\int_{\mathbb{R}^n} f d\gamma_n = 0 \text{ and } \int_{\mathbb{R}^n} f^2 d\gamma_n = 1.$$

We apply the logarithmic Sobolev inequality for $1+\varepsilon f$ where $\varepsilon > 0$ is small: we have

$$\int_{\mathbb{R}^n} (1+\varepsilon f)^2 \log(1+\varepsilon f) d\gamma_n - \frac{1}{2}\int_{\mathbb{R}^n} (1+\varepsilon f)^2 \log\left(\int_{\mathbb{R}^n} (1+\varepsilon f)^2\right) d\gamma_n$$
$$\leq \varepsilon^2 \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n,$$

which implies

$$\int_{\mathbb{R}^n} \{\varepsilon f - \frac{\varepsilon^2 f^2}{2} + 2\varepsilon^2 f^2 + O(\varepsilon^3)\} d\gamma_n - \frac{1}{2}\int_{\mathbb{R}^n} (1+2\varepsilon f + \varepsilon^2 f^2)$$
$$\times \log\left(1 + 2\varepsilon \int_{\mathbb{R}^n} f + \varepsilon^2 \int_{\mathbb{R}^n} f^2\right) d\gamma_n \leq \varepsilon^2 \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n.$$

Using the assumptions about f, we arrive at

$$1 \leq \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n + O(\varepsilon).$$

Letting $\varepsilon \to 0$ we see that

$$\int_{\mathbb{R}^n} f^2 d\gamma_n = 1 \leq \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n.$$

This completes the proof. $\square$

Our second application is the concentration inequality for Lipschitz functions (this is equivalent to Theorem 1.4.2).

Theorem 1.4.5 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz function with $\|f\|_{Lip} \leq 1$. Then,*

$$\gamma\left(x : \left|f(x) - \int_{\mathbb{R}^n} f d\gamma\right| > r\right) \leq 2e^{-r^2/2}.$$

Theorem 1.4.5 is a consequence of the following Proposition.

Proposition 1.4.2 *If $F : \mathbb{R}^n \to \mathbb{R}$ is a Lipschitz function with $\|F\|_{Lip} \le 1$ and if $\int_{\mathbb{R}^n} F d\gamma_n = 0$, then, for every $\lambda > 0$*

$$\int_{\mathbb{R}^n} e^{\lambda F} d\gamma \le e^{\lambda^2/2}.$$

Proof: Assume that $\int_{\mathbb{R}^n} F = 0$ and $\|F\|_{Lip} \le 1$. We set $f^2 = e^{\lambda F}$. Then,

$$\nabla f = \frac{\lambda e^{\lambda F} \nabla F}{2f} = \frac{\lambda e^{\lambda F} \nabla F}{2e^{\lambda F/2}} = \frac{\lambda}{2} e^{\lambda F/2} \nabla F.$$

Applying the logarithmic Sobolev inequality to f we have

$$\int_{\mathbb{R}^n} e^{\lambda F} \frac{\lambda F}{2} d\gamma - \frac{1}{2} \int_{\mathbb{R}^n} e^{\lambda F} d\gamma \log \left(\int_{\mathbb{R}^n} e^{\lambda F} d\gamma \right) \le \int_{\mathbb{R}^n} \frac{\lambda^2}{4} e^{\lambda F} |\nabla F|^2 d\gamma.$$

Since $|\nabla F| \le 1$, we have

$$\int_{\mathbb{R}^n} e^{\lambda F} |\nabla F|^2 d\gamma \le \int_{\mathbb{R}^n} e^{\lambda F} d\gamma.$$

Therefore, if we define

$$H(\lambda) = \int_{\mathbb{R}^n} e^{\lambda F} d\gamma$$

we have

$$\frac{\lambda}{2} H'(\lambda) - \frac{1}{2} H(\lambda) \log H(\lambda) \le \frac{\lambda^2}{4} H(\lambda).$$

In other words, for the function $k(\lambda) = \frac{\log H(\lambda)}{\lambda}$ we have

$$k'(\lambda) = \frac{\lambda H'(\lambda) - H(\lambda) \log H(\lambda)}{\lambda^2 H(\lambda)} \le \frac{1}{2}.$$

Note that

$$\lim_{\lambda \to 0} \frac{\log H(\lambda)}{\lambda} = \lim_{\lambda \to 0} \frac{H'(\lambda)}{H(\lambda)} = \int_{\mathbb{R}^n} F d\gamma = 0.$$

It follows that

$$k(\lambda) \le k(0) + \frac{1}{2}\lambda = \frac{\lambda}{2},$$

which implies

$$\int_{\mathbb{R}^n} e^{\lambda F} d\gamma \le e^{\lambda^2/2}. \qquad \square$$

Proof of Theorem 1.4.5: Let f be a Lipschitz function with $\|f\|_{Lip} \le 1$. If we set $F = f - \int f d\gamma$, then $\int F d\gamma = 0$ and

$$|F(x) - F(y)| = |f(x) - f(y)| \le \|f\|_{Lip} |x - y| \le |x - y|.$$

That is, $\|F\|_{Lip} \leq 1$. Proposition 1.4.1 shows that

$$\int_{\mathbb{R}^n} e^{\lambda F} d\gamma \leq e^{\lambda^2/2}.$$

Therefore,

$$e^{\lambda r}\gamma(\{x : F(x) \geq r\}) \leq \int_{\{F\geq r\}} e^{\lambda F} d\gamma \leq \int_{\mathbb{R}^n} e^{\lambda F} d\gamma \leq e^{\lambda^2/2}.$$

Thus, for every $\lambda > 0$

$$\gamma\left(\left\{x : f(x) - \int_{\mathbb{R}^n} f d\gamma \geq r\right\}\right) \leq e^{\frac{\lambda^2}{2}-\lambda r}.$$

Choosing $\lambda = r$ we get

$$\gamma\left(\left\{x : f(x) - \int f d\gamma \geq r\right\}\right) \leq e^{-r^2/2}.$$

Applying the same reasoning for $-f$ we conclude that

$$\gamma\left(x : \left|f(x) - \int f d\gamma\right| \geq r\right) \leq 2e^{-r^2/2}. \qquad \square$$

1.5 Applications of the Brunn-Minkowski inequality

The Brunn-Minkowski inequality is of fundamental importance. In this section we present a few of its applications, the ones which are more closely related to the topics we discuss in these notes.

1.5.1 An inequality of Rogers and Shephard

The **difference body** of a convex body K is the symmetric convex body

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

From the Brunn-Minkowski inequality it is clear that $|K-K| \geq 2^n|K|$ 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 1.5.1 *Let K be a convex body in $\mathbb{R}^n$. Then,*

$$|K - K| \leq \binom{2n}{n}|K|.$$

Proof: The Brunn-Minkowski inequality enters the proof through the observation that $f(x) = |K \cap (x+K)|^{1/n}$ is a concave function supported on $K-K$.

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 $0 \leq r \leq \rho_{K-K}(\theta)$. [By ρ_W we denote the **radial function** of W

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

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,

$$\begin{aligned}
\int_{K-K} |K \cap (x+K)| dx &= \int_{K-K} f^n(x) dx \geq \int_{K-K} g^n(x) dx \\
&= [f(0)]^n n\omega_n \int_{S^{n-1}} \int_0^{\rho_{K-K}(\theta)} r^{n-1}(1 - r/\rho)^n dr \sigma(d\theta) \\
&= n\omega_n |K| \int_{S^{n-1}} \rho_{K-K}^n(\theta) \sigma(d\theta) \int_0^1 t^{n-1}(1-t)^n dt \\
&= |K||K-K| \frac{n\Gamma(n)\Gamma(n+1)}{\Gamma(2n+1)} \\
&= \binom{2n}{n}^{-1} |K||K-K|.
\end{aligned}$$

On the other hand, Fubini's theorem gives

$$\begin{aligned}
\int_{K-K} |K \cap (x+K)| dx &= \int_{\mathbb{R}^n} |K \cap (x+K)| dx \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \chi_K(y) \chi_{x+K}(y) dy dx \\
&= \int_{\mathbb{R}^n} \chi_K(y) \left(\int_{\mathbb{R}^n} \chi_{y-K}(x) dx \right) dy \\
&= \int_K |y-K| dy = |K|^2.
\end{aligned}$$

Combining the above, we conclude the proof. □

Remark: 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 1.5.1 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 1.5.1 rests upon the fact that the volume of $K-K$ is not much larger than the volume of K. We have

$$|K-K|^{1/n} \leq 4|K|^{1/n},$$

which means that every convex body (which contains the origin) is contained in a symmetric convex body with the same more or less volume radius. This observation will find many applications in the sequel.

1.5.2 Hyperplane sections of a convex body

Consider a convex body K in $\mathbb{R}^n$, and fix a direction $\theta \in S^{n-1}$. We define a function $f = f_{K,\theta} : \mathbb{R} \to \mathbb{R}^+$ by

$$f(t) = |K \cap (\theta^\perp + t\theta)|.$$

Here $|\cdot|$ denotes $(n-1)$-dimensional volume. So, $f(t)$ gives the area of the section of K with the hyperplane which is perpendicular to θ, at a distance t from $\theta^\perp$.

Theorem 1.5.2 *Let K be a convex body in $\mathbb{R}^n$, $\theta \in S^{n-1}$, and*

$$f(t) = |K \cap (\theta^\perp + t\theta)|.$$

Then, $f^{1/(n-1)}$ is concave on its support.

Proof: We may assume that $\theta = e_n$, and identify $\theta^\perp$ with $\mathbb{R}^{n-1}$. For every t we set

$$K(t) = \{x \in \mathbb{R}^{n-1} \mid (x,t) \in K\}.$$

Let $I = \{t \mid K(t) \neq \emptyset\}$. Then, $K(t)$ is convex for every $t \in I$, and if $t, s \in I$, $\lambda \in (0,1)$, then

$$\lambda K(t) + (1-\lambda)K(s) \subseteq K(\lambda t + (1-\lambda)s).$$

Applying the Brunn-Minkowski inequality on $\mathbb{R}^{n-1}$, we get

$$|K(\lambda t + (1-\lambda)s)|^{\frac{1}{n-1}} \geq \lambda |K(t)|^{\frac{1}{n-1}} + (1-\lambda)|K(s)|^{\frac{1}{n-1}}.$$

Since $f(t) = |K \cap (\theta^\perp + t\theta)| = |K(t)|$, the proof is complete. □

Remark As we saw in Section 1.2, one can prove Theorem 1.5.2 by Schwarz symmetrization and then deduce the Brunn-Minkowski inequality from it.

Corollary 1.5.1 *With the notation of Theorem 1.5.2, f is a log-concave function on its support.* □

Corollary 1.5.2 *If K is a symmetric convex body, then $\|f\|_\infty = f(0)$.*

Proof: Since K is symmetric, we have $K(-t) = -K(t)$ for every $t \in I$. Thus, f is an even, log-concave function on I. It follows that

$$f(0) = f\left(\frac{t+(-t)}{2}\right) \geq \sqrt{f(t)}\sqrt{f(-t)} = f(t)$$

for every $t \in I$. $\square$

If we don't assume symmetry of K, we still have that the hyperplane section passing through the centroid of K is comparable to the maximal section in this direction.

Proposition 1.5.1 *Let K be a convex body with volume 1 and centroid at 0. If $\theta \in S^{n-1}$ and $f(t) = |K \cap (\theta^{\perp} + t\theta)|$, then*

$$\|f\|_{\infty} \leq ef(0) = e|K \cap \theta^{\perp}|.$$

Proof: Let $[-a, b]$ be the support of f, where $a, b > 0$, and assume that $\|f\|_{\infty} = f(t_0)$. We will show that

$$f(0) \geq \left(\frac{n}{n+1}\right)^{n-1} f(t_0) \geq \frac{1}{e} f(t_0).$$

Replacing θ by $-\theta$ if needed, we may assume that $0 < t_0 \leq b$. We can also assume that $f(0) < f(t_0)$ (otherwise, there is nothing to prove). Since K has its centroid at 0, we have

$$\int_{-a}^{b} t[f(t)]dt = \int_K \langle x, \theta\rangle dx = 0.$$

Therefore,

$$\int_{-a}^{0} (-t)[f(t)]dt = \int_0^b t[f(t)]dt \geq \int_0^{t_0} t[f(t)]dt.$$

We set $h = f^{\frac{1}{n-1}}$ and consider the linear function g defined by the equations $g(0) = h(0)$ and $g(t_0) = h(t_0)$. By the Brunn-Minkowski inequality, h is concave, which implies that $g \geq h$ on $[-a, 0]$ and $g \leq h$ on $[0, t_0]$. Since $g(0) < g(t_0)$, we see that g is strictly increasing and $g(-\gamma) = 0$ for some $-\gamma < -a$. In other words, $g(t) = c(t + \gamma)$ for some $c > 0$. Then,

$$\begin{aligned}
\int_{-\gamma}^{0} (-t)[g(t)]^{n-1}dt &\geq \int_{-a}^{0} (-t)[g(t)]^{n-1}dt \geq \int_{-a}^{0} (-t)[h(t)]^{n-1}dt \\
&\geq \int_0^{t_0} t[h(t)]^{n-1}dt \geq \int_0^{t_0} t[g(t)]^{n-1}dt.
\end{aligned}$$

This means that

$$\begin{aligned}
0 &\geq \int_{-\gamma}^{t_0} t[g(t)]^{n-1}dt = c^{n-1}\int_{-\gamma}^{t_0} t(t+\gamma)^{n-1}dt \\
&= c^{n-1}\int_0^{\gamma+t_0} t^{n-1}(t-\gamma)dt = c^{n-1}\left(\frac{(\gamma+t_0)^{n+1}}{n+1} - \gamma\frac{(\gamma+t_0)^n}{n}\right),
\end{aligned}$$

which implies

$$t_0 \leq \frac{\gamma}{n}.$$

Therefore,

$$\begin{aligned} \frac{f(0)}{f(t_0)} &= \left(\frac{g(0)}{g(t_0)}\right)^{n-1} \\ &= \left(\frac{\gamma}{\gamma + t_0}\right)^{n-1} \geq \left(\frac{n}{n+1}\right)^{n-1}. \quad \square \end{aligned}$$

1.5.3 Borell's lemma

Theorem 1.5.3 *Let K be a convex body in $\mathbb{R}^n$ with volume $|K| = 1$, and let A be a closed, convex and symmetric set such that $|K \cap A| = \delta > \frac{1}{2}$. Then, for every $t > 1$ we have*

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

Proof: We first show that

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

If this is not so, there exists $a \in A$ which can be written in the form

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

for some $a_1 \in A$ and $y \notin tA$. Then,

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

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

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

This proves the theorem. $\square$

Borell's lemma implies **concentration of volume** in $\mathbb{R}^n$: If $A \cap K$ captures more than half of the volume of K, then the percentage of K which stays outside tA, $t > 1$ decreases exponentially with respect to t as $t \to \infty$, at a rate which is independent from the body K and the dimension n. We shall give a first application of this fact, showing that linear functionals on convex bodies satisfy inverse Hölder inequalities.

Theorem 1.5.4 *There exists an absolute constant $c > 0$ with the following property. If K is a convex body in $\mathbb{R}^n$ with volume $|K| = 1$, and $\theta \in \mathbb{R}^n$, then for every $p > 1$,*

$$\left(\int_K |\langle x, \theta\rangle|^p dx\right)^{\frac{1}{p}} \leq cp \int_K |\langle x, \theta\rangle| dx.$$

Proof: We define $I_1 = \int_K |\langle x, \theta\rangle| dx$ and set

$$A = \{x \in \mathbb{R}^n : |\langle x, \theta\rangle| \leq 3I_1\}.$$

Now, A is obviously convex and symmetric, and

$$I_1 = \int_K |\langle x, \theta\rangle| dx \geq \int_{K\cap A^c} |\langle x, \theta\rangle| dx \geq 3I_1 |K \cap A^c|,$$

which shows that

$$|A \cap K| \geq \frac{2}{3}.$$

We write

$$\int_K |\langle x, \theta\rangle|^p dx = \int_0^\infty pt^{p-1} |\{x \in K : |\langle x, \theta\rangle| > t\}| dt,$$

and observe that $\{x \in K : |\langle x, \theta\rangle| > t\} = K \cap (\mathbb{R}^n \backslash (t/3I_1)A)$. So, Borell's lemma applies for $t > 3I_1$. We write

$$\begin{aligned}\int_K |\langle x, \theta\rangle|^p dx \quad &= \quad \int_0^{3I_1} pt^{p-1} |K \cap (\mathbb{R}^n \backslash (t/3I_1)A)| dt \\ &\quad + \int_{3I_1}^\infty pt^{p-1} |K \cap (\mathbb{R}^n \backslash (t/3I_1)A)| dt.\end{aligned}$$

The first integral is simply bounded by

$$\int_0^{3I_1} pt^{p-1} dt = (3I_1)^p,$$

while, after the substitution $t = 3I_1 s$, the second one becomes

$$(3I_1)^p \int_1^\infty ps^{p-1} |K \cap (\mathbb{R}^n \backslash sA)| ds \leq (3I_1)^p \int_1^\infty ps^{p-1} 2^{-s/2} ds.$$

A simple computation with the Gamma-function gives

$$\int_K |\langle x, \theta\rangle|^p dx \leq (3I_1)^p (c_1 p)^p$$

for some absolute constant $c_1 > 0$, which gives

$$\left(\int_K |\langle x, \theta\rangle|^p dx\right)^{\frac{1}{p}} \leq cpI_1 = cp \int_K |\langle x, \theta\rangle| dx. \qquad \square$$

1.5.4 The isoperimetric inequality for convex bodies

The **surface area** (or Minkowski content) $A(K)$ of a convex body K is defined by

$$A(K)=\lim_{t\to 0^+}\frac{|K+tD_n|-|K|}{t}.$$

It is a well-known fact that among all convex bodies of a given volume the ball has minimal surface area. This is an immediate consequence of the Brunn-Minkowski inequality: If K is a convex body in $\mathbb{R}^n$ with $|K|=|rD_n|$, then for every $t>0$

$$|K+tD_n|^{1/n}\geq |K|^{1/n}+t|D_n|^{1/n}=(r+t)|D_n|^{1/n}.$$

It follows that the surface area $A(K)$ of K satisfies

$$A(K)=\lim_{t\to 0^+}\frac{|K+tD_n|-|K|}{t}\geq \lim_{t\to 0^+}\frac{(r+t)^n-r^n}{t}|D_n|=n|D_n|^{\frac{1}{n}}|K|^{\frac{n-1}{n}}$$

with equality if $K=rD_n$. The question of uniqueness in the equality case is more delicate.

Note that a stronger isoperimetric statement holds true: If, say, $|K|=|D_n|$ then

$$|K+tD_n|\geq |D_n+tD_n|$$

for every $t>0$. In other words, for fixed volume and for every $t>0$, the t-extension

$$K_t=\{y\mid d(y,K)\leq t\}$$

has minimal volume if K is a ball.

1.5.5 The spherical isoperimetric inequality

Consider the unit sphere S^{n-1} with the geodesic distance ρ and the rotationally invariant probability measure σ. For every Borel subset A of S^{n-1} and for every $t>0$, we define the t-extension of A:

$$A_t=\{x\in S^{n-1}:\rho(x,A)\leq t\}.$$

Then, the isoperimetric inequality for the sphere is the following statement:

> "Among all Borel subsets A of S^{n-1} with given measure $\alpha\in(0,1)$, a spherical cap $B(x,r)$ of radius $r>0$ such that $\sigma(B(x,r))=\alpha$ has minimal t-extension for every $t>0$."

This means that if $A\subseteq S^{n-1}$ and $\sigma(A)=\sigma(B(x_0,r))$ for some $x_0\in S^{n-1}$ and $r>0$, then

$$\sigma(A_t)\geq \sigma(B(x_0,r+t))$$

for every $t>0$. Since the σ-measure of a cap is easily computable, one can give a lower bound for the measure of the t-extension of any subset of the sphere. We are mainly interested in the case $\sigma(A)=\frac{1}{2}$, and a straightforward computation shows the following:

Theorem 1.5.5 *If A is a Borel subset of S^{n+1} and $\sigma(A) = 1/2$, then*

$$\sigma(A_t) \geq 1 - \sqrt{\pi/8}\exp(-t^2 n/2)$$

for every $t > 0$. $\square$

[The constant $\sqrt{\pi/8}$ may be replaced by a sequence of constants a_n tending to $\frac{1}{2}$ as $n \to \infty$.]

The spherical isoperimetric inequality can be proved by spherical symmetrization techniques. However, it was recently observed that one can give a very simple proof of an analogous exponential estimate using the Brunn-Minkowski inequality. The key point is the following lemma:

Lemma 1.5.1 *Consider the probability measure $\mu(A) = |A|/|D_n|$ on the Euclidean unit ball D_n. If A, B are subsets of D_n with $\mu(A) \geq \alpha$, $\mu(B) \geq \alpha$, and if $\rho(A,B) = \inf\{|a-b| : a \in A, b \in B\} = \rho > 0$, then*

$$\alpha \leq \exp(-\rho^2 n/8).$$

[In other words, if two disjoint subsets of D_n have positive distance ρ, then at least one of them must have small volume (depending on ρ) when the dimension n is high.]

Proof: We may assume that A and B are closed. By the Brunn-Minkowski inequality, $\mu(\frac{A+B}{2}) \geq \alpha$. On the other hand, the parallelogram law shows that if $a \in A, b \in B$ then

$$|a+b|^2 = 2|a|^2 + 2|b|^2 - |a-b|^2 \leq 4 - \rho^2.$$

It follows that $\frac{A+B}{2} \subseteq \sqrt{1 - \frac{\rho^2}{4}}D_n$, hence

$$\mu\left(\frac{A+B}{2}\right) \leq \left(1 - \frac{\rho^2}{4}\right)^{n/2} \leq \exp(-\rho^2 n/8). \quad \square$$

Proof of Theorem 1.5.5 (with weaker constants). Assume that $A \subseteq S^{n-1}$ with $\sigma(A) = 1/2$. Let $t > 0$ and define $B = (A_t)^c \subseteq S^{n-1}$. We fix $\lambda \in (0,1)$ and consider the subsets

$$\tilde{A} = \bigcup\{tA : \lambda \leq t \leq 1\} \quad \text{and} \quad \tilde{B} = \bigcup\{tB : \lambda \leq t \leq 1\}$$

of D_n. These are disjoint with distance $\simeq \lambda t$. Lemma 1.5.1 shows that $\mu(\tilde{B}) \leq \exp(-c\lambda^2 t^2 n/8)$, and since $\mu(\tilde{B}) = (1-\lambda^n)\sigma(B)$ we obtain

$$\sigma(A_t) \geq 1 - \frac{1}{1-\lambda^n}\exp(-c\lambda^2 t^2 n/8).$$

We conclude the proof by choosing $\lambda = 1/2$. $\square$

1.5.6 Blaschke-Santaló inequality

Let K be a symmetric convex body in $\mathbb{R}^n$. The **polar body** of K is the symmetric convex body

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

The **volume product** $s(K)$ of K is the product $|K| \cdot |K^\circ|$ of the volumes of K and its polar. It is easily seen that the volume product is an invariant of the class of K: If $T \in GL(n)$, then

$$s(K) = |K| \cdot |K^\circ| = |TK| \cdot |(TK)^\circ| = s(TK).$$

The Blaschke-Santaló inequality asserts that $s(K)$ is maximized when K is an ellipsoid.

Theorem 1.5.6 *Let K be a symmetric convex body in $\mathbb{R}^n$. Then, $|K| \cdot |K^\circ| \le \omega_n^2$.*

We will give a proof of this fact which is based on Steiner symmetrization.

Theorem 1.5.7 *Let K be a symmetric convex body in $\mathbb{R}^n$ and let $\theta \in S^{n-1}$. If $K_1 = S_\theta(K)$ is the Steiner symmetrization of K in the direction of θ, then*

$$|K||K^\circ| \le |K_1||(K_1)^\circ|.$$

Proof: Steiner symmetrization preserves volume. Therefore, it suffices to prove that $|K^\circ| \le |(K_1)^\circ|$.

For simplicity we assume that $\theta^\perp = \mathbb{R}^{n-1}$. It is not hard to see that

$$S_\theta(K) = K_1 = \left\{(x, \frac{t_1 - t_2}{2}) : x \in P_{\theta^\perp} K, (x, t_1) \in K, (x, t_2) \in K\right\}.$$

As usual, for every $A \subseteq \mathbb{R}^n$ we write

$$A(t) = \{x \in \mathbb{R}^{n-1} : (x, t) \in A\}.$$

Claim: For every $s \in \mathbb{R}$,

$$\frac{K^\circ(s) + K^\circ(-s)}{2} \subseteq (K_1)^\circ(s).$$

Proof of the claim: Let $y_1 \in K^\circ(s)$ and $y_2 \in K^\circ(-s)$. Then, $(y_1, s) \in K^\circ$ and $(y_2, -s) \in K^\circ$. We want to show that $(y_1 + y_2)/2 \in (K_1)^\circ(s)$, or equivalently, $(\frac{y_1+y_2}{2}, s) \in (K_1)^\circ$.

Let $(x, \frac{t_1-t_2}{2}) \in K_1$, where $(x, t_1) \in K$ and $(x, t_2) \in K$. We have to check that

$$\langle (x, \frac{t_1 - t_2}{2}), (\frac{y_1 + y_2}{2}, s)\rangle \le 1.$$

But this last quantity is equal to

$$\langle x, \frac{y_1+y_2}{2}\rangle + \frac{st_1 - st_2}{2} = \frac{\langle (x,t_1),(y_1,s)\rangle + \langle (x,t_2),(y_2,-s)\rangle}{2} \leq 1. \qquad \square$$

By the Brunn-Minkowski inequality,

$$|(K_1)^\circ(s)| \geq |K^\circ(s)|^{\frac{1}{2}} |K^\circ(-s)|^{\frac{1}{2}},$$

and since $|K^\circ(s)| = |K^\circ(-s)|$ by the symmetry of K°, we get

$$|(K_1)^\circ(s)| \geq |K^\circ(s)|$$

for every $s \in \mathbb{R}$. Integration shows that

$$|(K_1)^\circ| = \int_{-\infty}^{+\infty} |(K_1)^\circ(s)| ds \geq \int_{-\infty}^{+\infty} |K^\circ(s)| ds = |K^\circ|. \qquad \square$$

The Blaschke-Santaló inequality now follows by applying a suitable sequence of Steiner symmetrizations to K.

1.5.7 Urysohn's inequality

Recall the definition of the support function h_K and of the width function of a convex body K. For every $u \in S^{n-1}$ we define

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

and

$$w_K(u) = h_K(u) + h_K(-u).$$

The **mean width** $w(K)$ of K is the average

$$w(K) = \int_{S^{n-1}} w_K(u)\sigma(du) = 2\int_{S^{n-1}} h_K(u)\sigma(du).$$

A classical inequality of Urysohn states that for fixed volume, Euclidean ball has minimal mean width.

Theorem 1.5.8 *Let K be a convex body in $\mathbb{R}^n$. Then,*

$$\frac{w(K)}{2} \geq \left(\frac{|K|}{|D_n|}\right)^{1/n}.$$

We will give a simple proof of this fact using Steiner symmetrization. By continuity of the mean width with respect to the Hausdorff metric, it suffices to prove the following.

Proposition 1.5.2 *Let K be a convex body in $\mathbb{R}^n$ and let $\theta \in S^{n-1}$. Then,*

$$w(S_\theta K) \leq w(K).$$

Proof: We may assume that $\theta = e_n$ and write

$$S_\theta(K) = K_1 = \left\{ (x, \frac{t_1 - t_2}{2}) : x \in P_{\theta^\perp} K, (x, t_1) \in K, (x, t_2) \in K \right\}.$$

Let $u = (u_1, \ldots, u_n) \in S^{n-1}$. We set $u' = (u_1, \ldots, u_{n-1}, -u_n)$. Then,

$$\begin{aligned} h_{S_\theta K}(u) &= \max \left\{ \langle (x, \frac{t_1 - t_2}{2}), u \rangle \mid (x, t_i) \in K \right\} \\ &= \frac{1}{2} \max\{\langle (x, t_1), u \rangle \mid (x, t_1) \in K\} \\ &\quad + \frac{1}{2} \max\{\langle (x, -t_2), u \rangle \mid (x, t_2) \in K\} \\ &= \frac{1}{2} \max\{\langle (x, t_1), u \rangle \mid (x, t_1) \in K\} \\ &\quad + \frac{1}{2} \max\{\langle (x, t_2), u' \rangle \mid (x, t_2) \in K\} \\ &= \frac{1}{2} h_K(u) + \frac{1}{2} h_K(u'). \end{aligned}$$

Since u and u' have the same distribution in S^{n-1}, we see that

$$\begin{aligned} w(S_\theta K) &= 2 \int_{S^{n-1}} h_{S_\theta K}(u) \sigma(du) \\ &\leq \int_{S^{n-1}} h_K(u) \sigma(du) + \int_{S^{n-1}} h_K(u') \sigma(du) \\ &= w(K). \qquad \square \end{aligned}$$

Chapter 2

Rearrangement inequalities

2.1 Introduction

The purpose of this Chapter is to present the Brascamp-Lieb inequality and its reverse form. The Brascamp-Lieb inequality concerns the multilinear operator $\Phi : L^{p_1}(\mathbb{R}) \times \cdots \times L^{p_m}(\mathbb{R}) \to \mathbb{R}$ defined by

$$\Phi(f_1, \dots, f_m) = \int_{\mathbb{R}^n} \prod_{j=1}^{m} f_j(\langle u_j, x\rangle)\, dx, \tag{1.1}$$

where $m \geq n$, $p_1, \dots, p_m \geq 1$ with $\frac{1}{p_1} + \cdots + \frac{1}{p_m} = n$, and $u_1, \dots, u_m \in \mathbb{R}^n$.

Brascamp and Lieb proved that the norm of Φ is the supremum of

$$\frac{\Phi(g_1, \dots, g_m)}{\prod_{j=1}^{m} \|g_j\|_{p_j}} \tag{1.2}$$

over all centered Gaussian functions $g_1, \dots, g_m$, i.e. over all functions of the form

$$g_j(t) = e^{-\lambda_j t^2}, \ \ \lambda_j > 0. \tag{1.3}$$

Many sharp geometric inequalities that we are going to see later on make heavy use of this fact.

The original proof of the Brascamp-Lieb inequality was based on a general rearrangement inequality of Brascamp, Lieb and Luttinger, which is the subject of Section 2.2. We introduce the symmetric decreasing rearrangement f^* of a Borel measurable function f vanishing at infinity and show that

$$\Phi(f_1, \dots, f_m) \leq \Phi(f_1^*, \dots, f_m^*). \tag{1.4}$$

This reduces the problem to the class of radially symmetric and non-increasing functions vanishing at infinity. A generalization to functions of several variables is also obtained by Steiner symmetrization: this was actually needed for the original proof of the Brascamp-Lieb inequality.

In Section 2.3 we go back to the Brunn-Minkowski inequality and discuss an inequality of Szarek and Voiculescu about the volume of restricted Minkowski sums. A restricted sum of two subsets A and B of $\mathbb{R}^n$ is a set of the form

$$A +_\Theta B = \{x + y \mid (x, y) \in \Theta\}, \tag{1.5}$$

where Θ is a subset of $A \times B$. Szarek and Voiculescu proved the following. If $\rho \in (0, 1)$, $A, B \subset \mathbb{R}^n$ with

$$|B|^{1/n} = \rho |A|^{1/n} \tag{1.6}$$

and Θ is a subset of $A \times B$ with

$$|\Theta| \geq (1 - c \min\{\rho\sqrt{n}, 1\})|A| \cdot |B|, \tag{1.7}$$

then,

$$|A +_\Theta B|^{2/n} \geq |A|^{2/n} + |B|^{2/n}. \tag{1.8}$$

The proof of this inequality requires the multivariable rearrangement inequality of Brascamp, Lieb and Luttinger. As an application, we give a proof of Shannon's entropy power inequality through restricted Minkowski sums.

Section 2.4 is devoted to the argument of Brascamp and Lieb. In Section 2.5 we discuss the reverse (or dual) form of the Brascamp-Lieb inequality which was conjectured by Ball in connection with problems of Convex Geometric Analysis and was proved by Barthe. Gaussian functions are the extremal functions in this inequality as well. Barthe's argument exploits the transportation of measure idea which was already used in the proof of the Prékopa-Leindler inequality. It gives a new proof of the Brascamp-Lieb inequality at the same time.

Finally, in Section 2.6 we discuss the multidimensional version of the Brascamp-Lieb inequality and its reverse form. It concerns the multilinear operator $\Psi : L^{p_1}(\mathbb{R}^{n_1}) \times \cdots \times L^{p_m}(\mathbb{R}^{n_m}) \to \mathbb{R}$ defined by

$$\Psi(f_1, \ldots, f_m) = \int_{\mathbb{R}^n} \prod_{j=1}^m f_j(B_j x) dx, \tag{1.9}$$

where B_j is a linear mapping from $\mathbb{R}^n$ onto $\mathbb{R}^{n_j}$, $p_j \geq 1$ and certain natural conditions on n_j, p_j and B_j are satisfied. Lieb proved that Gaussian functions are again the maximizers. We present Barthe's argument which proves a dual form of this inequality as well. Here, the main technical tool is the Brenier map which was discussed in Section 1.3.

2.2 A general rearrangement inequality

2.2.1 Symmetric decreasing rearrangement

In this Section we consider Borel measurable functions $f : \mathbb{R}^n \to \mathbb{R}^+$ which satisfy the following: for every $t > 0$, the set $\{x \in \mathbb{R}^n : f(x) > t\}$ has finite Lebesgue measure. We say that such an f **vanishes at infinity**.

Let A be a Borel subset of $\mathbb{R}^n$ with finite Lebesgue measure. The **symmetric rearrangement** A^* of A is the open ball with centre at the origin, whose volume is equal to the measure of A. Since we choose A^* to be open, χ_{A^*} is lower semicontinuous.

The **symmetric decreasing rearrangement** of χ_A is defined by

$$\chi_A^* = \chi_{A^*}. \tag{2.1}$$

If $f : \mathbb{R}^n \to \mathbb{R}^+$ is a Borel measurable function which vanishes at infinity, we set

$$f^*(x) = \int_0^\infty \chi^*_{\{f>t\}}(x)dt = \int_0^\infty \chi_{\{f>t\}^*}(x)dt. \tag{2.2}$$

Several basic properties of the symmetric decreasing rearrangement follow.

1. f^* is a non-negative function.
2. f^* is radially symmetric and non-increasing:

   $$f^*(x) = f^*(y) \quad \text{if} \quad |x| = |y| \tag{2.3}$$

   and

   $$f^*(x) \geq f^*(y) \quad \text{if} \quad |x| \leq |y|. \tag{2.4}$$

3. f^* is lower semicontinuous: for every $t > 0$ we have $\{x : f^*(x) > t\} = \{x : f(x) > t\}^*$, which is an open set. In particular, f^* is measurable.
4. The level sets of f^* are the symmetric rearrangements of the corresponding level sets of f: for every $t > 0$,

   $$\{x : f^*(x) > t\} = \{x : f(x) > t\}^*. \tag{2.5}$$

5. If f and g are non-negative Borel measurable functions on $\mathbb{R}^n$ which vanish at infinity, and if $f(x) \leq g(x)$ for all $x \in \mathbb{R}^n$, then $f^*(x) \leq g^*(x)$ for all $x \in \mathbb{R}^n$. This follows from the fact that the level sets of g contain the corresponding level sets of f.
6. We easily check that

   $$\|f\|_p = \|f^*\|_p \tag{2.6}$$

for all $1 \leq p \leq \infty$. For example, if $1 \leq p < \infty$, using (2.5) we have

$$\begin{aligned} \|f\|_p^p &= p \int_0^\infty t^{p-1} |\{x : f(x) > t\}| dt \\ &= p \int_0^\infty t^{p-1} |\{x : f(x) > t\}^*| dt \\ &= p \int_0^\infty t^{p-1} |\{x : f^*(x) > t\}| dt \\ &= \|f^*\|_p^p. \end{aligned}$$

There are many well-known inequalities about rearrangements. The simplest one is the following.

Theorem 2.2.1 *Let $f, g : \mathbb{R}^n \to \mathbb{R}^+$ be two non-negative Borel measurable functions vanishing at infinity. Then,*

$$\int_{\mathbb{R}^n} f(x)g(x)dx \leq \int_{\mathbb{R}^n} f^*(x)g^*(x)dx. \tag{2.7}$$

Proof: We write

$$\int_{\mathbb{R}^n} f(x)g(x)dx = \int_0^\infty \int_0^\infty \int_{\mathbb{R}^n} \chi_{\{f>t\}}(x)\chi_{\{g>s\}}(x)dx\ ds\ dt$$

and

$$\int_{\mathbb{R}^n} f^*(x)g^*(x)dx = \int_0^\infty \int_0^\infty \int_{\mathbb{R}^n} \chi_{\{f^*>t\}}(x)\chi_{\{g^*>s\}}(x)dx\ ds\ dt.$$

Thus, it suffices to show that

$$|\{f > t\} \cap \{g > s\}| \leq |\{f^* > t\} \cap \{g^* > s\}| \tag{2.7}$$

for all $t, s > 0$. If we set $F_t = \{f > t\}$ and $G_s = \{g > s\}$, then (2.7) reduces to the inequality

$$|F_t \cap G_s| \leq |F_t^* \cap G_s^*|. \tag{2.8}$$

This is clearly true: F_t^* and G_s^* are centered balls, and hence one of them is contained in the other. It follows that

$$|F_t^* \cap G_s^*| = \min\{|F_t^*|, |G_s^*|\} = \min\{|F_t|, |G_s|\} \geq |F_t \cap G_s|. \qquad \square$$

2.2.2 Brascamp-Lieb-Luttinger inequality

A rearrangement inequality of Riesz states that if $f, g, h \in L^1(\mathbb{R})$, then

$$\int\int_{\mathbb{R}^2} f(x)g(y)h(x-y)\,dy\,dx \leq \int\int_{\mathbb{R}^2} f^*(x)g^*(y)h^*(x-y)\,dy\,dx. \tag{2.9}$$

Brascamp, Lieb and Luttinger generalized this inequality as follows.

Theorem 2.2.2 *Assume that $f_1, \dots, f_m : \mathbb{R} \to \mathbb{R}^+$ are measurable functions and $u_1, \dots, u_m \in \mathbb{R}^n$. Then,*

$$\int_{\mathbb{R}^n} \prod_{j=1}^m f_j(\langle x, u_j\rangle)dx \leq \int_{\mathbb{R}^n} \prod_{j=1}^m f_j^*(\langle x, u_j\rangle)dx. \tag{2.10}$$

Remark: It is clear that (2.9) is a special case of (2.10): we just set $f_1 = f$, $f_2 = g$, $f_3 = h$ and $u_1 = (1,0)$, $u_2 = (0,1)$, $u_3 = (1,-1) \in \mathbb{R}^2$.

The proof of Theorem 2.2.2 is based on Brunn's concavity principle.

Proposition 2.2.1 *The theorem holds true if each f_j is the characteristic function of a closed bounded interval.*

Proof: Assume that $I_j = [b_j - c_j, b_j + c_j]$ and $f_j = \chi_{I_j}$, $j = 1, \dots, m$. We define $f_j(\cdot|t)$ by

$$f_j(x|t) = f_j(x + b_jt) = \chi_{I_j - b_jt}(x),$$

where $I_j - b_jt = [(1-t)b_j - c_j, (1-t)b_j + c_j]$. We also define $G : [0,1] \to \mathbb{R}$ by

$$G(t) = \int_{\mathbb{R}^n} \prod_{j=1}^m f_j(\langle u_j, x\rangle|t)dx. \tag{2.11}$$

Since $f_j(x|1) = \chi_{[-c_j,c_j]}(x) = f_j^*(x)$, the proposition will follow if we show that $G(0) \leq G(1)$.

We will prove something stronger: G is increasing. To this end, we consider the symmetric convex set

$$K = \{\overline{x} = (x, x_{n+1}) \in \mathbb{R}^{n+1} : -c_j \leq \langle u_j, x\rangle - b_jx_{n+1} \leq c_j\}. \tag{2.12}$$

Then,

$$|K \cap (e_{n+1}^{\perp} + te_{n+1})| = |\{x \in \mathbb{R}^n : -c_j \leq \langle u_j, x\rangle - tb_j \leq c_j\}| = G(1-t). \tag{2.13}$$

By Brunn's concavity principle, G is a decreasing function of $1-t$. $\square$

Proposition 2.2.2 *The theorem holds true if each f_j is the characteristic function of a finite union of closed bounded intervals.*

Proof: Assume that f_j is the characteristic function of the union of n_j disjoint closed intervals: $f_j = \chi_{A_j}$, where $A_j = \bigcup_{k=1}^{n_j}[b_{kj} - c_{kj}, b_{kj} + c_{kj}]$ and $b_{kj} + c_{kj} < b_{k+1,j} - c_{k+1,j}$, $k = 1, \ldots, n_j - 1$.

We will work by induction on $N = \{n_1, \ldots, n_m\}$. We say that $M \prec N$ if $m_j \leq n_j$ for every j and there exists i for which $m_i < n_i$. Proposition 2.2.1 covers the case $N = \{1, 1, \ldots, 1\}$ and we assume that N is given and the proposition has been proved for every $M \prec N$.

For every $t \in [0,1]$ we define

$$f_{kj}(x|t) = \chi_{[b_{kj}(1-t)-c_{kj}, b_{kj}(1-t)+c_{kj}]}(x), \qquad j = 1, \ldots, m, \ \ k = 1, \ldots, n_j.$$

The new distances between "successive intervals" are

$$b_{k+1,j}(1-t) - c_{k+1,j} - b_{kj}(1-t) - c_{kj}$$

and hence, they are decreasing functions of t. We define

$$f_j(x,t) = \sum_{k=1}^{n_j} f_{kj}(x|t), \qquad j = 1, \ldots, m$$

and set

$$\tau = \min_{k,j}\left(1 - \frac{c_{kj} + c_{k+1,j}}{b_{k+1,j} - b_{kj}}\right).$$

Thus, τ is the smallest number t in $(0,1]$ for which two successive intervals of some $f_j(\cdot|t)$ will join. This means that if $M(\tau) = \{n_1(\tau), \ldots n_m(\tau)\}$ is the m-tuple which corresponds to $f_1(\cdot|\tau), \ldots, f_m(\cdot|\tau)$, we have $M(\tau) \prec N$.

We write

$$\begin{aligned}
\int_{\mathbb{R}^n} \prod_{j=1}^m f_j(\langle u_j, x\rangle)dx &= \int_{\mathbb{R}^n} \prod_{j=1}^m \sum_{k=1}^{n_j} f_{kj}(\langle u_j, x\rangle)dx \\
&= \sum_{k_1=1}^{n_1} \cdots \sum_{k_m=1}^{n_m} \int_{\mathbb{R}^n} \prod_{j=1}^m f_{k_j j}(\langle u_j, x\rangle)dx
\end{aligned}$$

where $f_{kj} = \chi_{[b_{kj}-c_{kj}, b_{kj}+c_{kj}]}$. The fact that the function G in the proof of Proposition 2.2.1 is increasing shows that

$$\int_{\mathbb{R}^n} \prod_{j=1}^m f_{k_j j}(\langle u_j, x\rangle)dx \leq \int_{\mathbb{R}^n} \prod_{j=1}^m f_{k_j j}(\langle u_j, x\rangle|\tau)dx \tag{2.14}$$

If we add and change the order of integration and summation, we obtain

$$\begin{aligned}
\int_{\mathbb{R}^n} \prod_{j=1}^m f_j(\langle u_j, x\rangle)dx &\leq \int_{\mathbb{R}^n} \prod_{j=1}^m \left(\sum_{k=1}^{n_j} f_{kj}(\langle u_j, x\rangle|\tau)\right) dx \\
&= \int_{\mathbb{R}^n} \prod_{j=1}^m f_j(\langle u_j, x\rangle|\tau)dx.
\end{aligned}$$

Now, we apply the inductive hypothesis for $f_j(\cdot|\tau)$. We have

$$\int_{\mathbb{R}^n}\prod_{j=1}^m f_j(\langle u_j,x\rangle|\tau)dx \le \int_{\mathbb{R}^n}\prod_{j=1}^m f_j^*(\langle u_j,x\rangle|\tau)dx,$$

where $f_j^*(\cdot|\tau)$ is the symmetric decreasing rearrangement of $f_j(\cdot|\tau)$. On observing that f_j and $f_j(\cdot|\tau)$ have the same symmetric decreasing rearrangement, we get

$$\int_{\mathbb{R}^n}\prod_{j=1}^m f_j(\langle x,u_j\rangle)dx \le \int_{\mathbb{R}^n}\prod_{j=1}^m f_j^*(\langle x,u_j\rangle)dx. \qquad \square$$

Standard approximation arguments give the following.

Proposition 2.2.3 *The theorem holds true if each f_j is the characteristic function of a measurable set A_j with $|A_j|<\infty$, $j=1,\dots,m$.* $\square$

We can now give the

Proof of Theorem 2.2.2: Assuming that $f_1,\dots,f_m:\mathbb{R}\to\mathbb{R}^+$ are measurable functions and using Proposition 2.2.3, we write

$$\begin{aligned}
\int_{\mathbb{R}^n}\prod_{j=1}^m f_j(\langle x,u_j\rangle)dx &= \int_{\mathbb{R}^n}\prod_{j=1}^m\int_0^\infty \chi_{\{f_j>t_j\}}(\langle u_j,x\rangle)dt_j dx\\
&= \int_0^\infty\cdots\int_0^\infty\left(\int_{\mathbb{R}^n}\prod_{j=1}^m \chi_{\{f_j>t_j\}}(\langle u_j,x\rangle)dx\right)dt_m\dots dt_1\\
&\le \int_0^\infty\cdots\int_0^\infty\left(\int_{\mathbb{R}^n}\prod_{j=1}^m \chi^*_{\{f_j>t_j\}}(\langle u_j,x\rangle)dx\right)dt_m\dots dt_1\\
&= \int_0^\infty\cdots\int_0^\infty\left(\int_{\mathbb{R}^n}\prod_{j=1}^m \chi_{\{f_j^*>t_j\}}(\langle u_j,x\rangle)dx\right)dt_m\dots dt_1\\
&= \dots\\
&= \int_{\mathbb{R}^n}\prod_{j=1}^m f_j^*(\langle x,u_j\rangle)dx. \qquad \square
\end{aligned}$$

2.2.3 Generalization to functions of several variables

Let $f:\mathbb{R}^k\to\mathbb{R}^+$ be a measurable function vanishing at infinity. We consider a $(k-1)$-dimensional subspace V of $\mathbb{R}^k$ and fix a coordinate system so that e_1 is the normal vector to V. The **Steiner symmetrization** $f^*(\cdot|V)$ of f with respect to V is defined as follows: If $x_2, \ldots, x_k \in \mathbb{R}$, we consider the function $h(t) = f(t, x_2, \ldots, x_k)$ and define

$$f^*(t, x_2, \ldots, x_k|V) = h^*(t).$$

It is easily checked that $\{x : f^*(x|V) > t\}$ is the Steiner symmetrization of $\{x : f(x) > t\}$ with respect to V. In particular, p-norms are preserved by $f \mapsto f^*(\cdot|V)$.

Lemma 2.2.1 *Let $f_1, \ldots, f_m : \mathbb{R}^k \to \mathbb{R}^+$ be measurable functions vanishing at infinity, and let A be an $m \times n$ matrix. If V is a $(k-1)$-dimensional subspace of $\mathbb{R}^k$, then*

$$\int_{\mathbb{R}^k} \cdots \int_{\mathbb{R}^k} \prod_{j=1}^m f_j\left(\sum_{i=1}^n a_{ji}x_i\right) dx_n \ldots dx_1$$

$$\leq \int_{\mathbb{R}^k} \cdots \int_{\mathbb{R}^k} \prod_{j=1}^m f_j^*\left(\sum_{i=1}^n a_{ji}x_i|V\right) dx_n \ldots dx_1.$$

Proof: We set $x_i = (t_i, y_i)$, $i = 1, \ldots, n$. Changing the order of integration, we get

$$\int_{\mathbb{R}^k} \cdots \int_{\mathbb{R}^k} \prod_{j=1}^m f_j\left(\sum_{i=1}^n a_{ji}x_i\right) dx_n \ldots dx_1$$

$$= \int_{\mathbb{R}^{k-1}} \cdots \int_{\mathbb{R}^{k-1}} \left(\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \prod_{j=1}^m f_j\left(\sum_{i=1}^n a_{ji}(t_i, y_i)\right) dt_n \ldots dt_1\right) dy_n \ldots dy_1$$

$$= \int_{\mathbb{R}^{k-1}} \cdots \int_{\mathbb{R}^{k-1}} \left(\int_{\mathbb{R}^n} \prod_{j=1}^m g_j(\langle u_j, t\rangle)dt\right) dy_n \ldots dy_1,$$

where $u_j = (u_{j1}, \ldots, u_{jn})$, $t = (t_1, \ldots, t_n)$ and, for fixed y_i,

$$g_j(\langle u_j, t\rangle) = f_j\left(\sum_{i=1}^n a_{ji}x_i\right).$$

Observe that

$$g_j^*(\langle u_j, t\rangle) = f_j^*\left(\sum_{i=1}^n a_{ji}x_i|V\right).$$

Then, Theorem 2.2.2 completes the proof. □

Successive symmetrizations with respect to $(k-1)$-dimensional subspaces yield the symmetric decreasing rearrangement f_j^* of each f_j, $j = 1, \ldots, m$. Thus, we have:

Theorem 2.2.3 *Let $f_1, \dots, f_m : \mathbb{R}^k \to \mathbb{R}^+$ be measurable functions vanishing at infinity, and let A be an $m \times n$ matrix. Then,*

$$\int_{\mathbb{R}^k} \cdots \int_{\mathbb{R}^k} \prod_{j=1}^m f_j(\sum_{i=1}^n a_{ji}x_i)dx_n \dots dx_1 \leq \int_{\mathbb{R}^k} \cdots \int_{\mathbb{R}^k} \prod_{j=1}^m f_j^*(\sum_{i=1}^n a_{ji}x_i)dx_n \dots dx_1 . \square$$

If we set $z_l = (x_{1l}, \dots, x_{nl})$, $l = 1, \dots, k$ and $u_j = (a_{j1}, \dots, a_{jn}) \in \mathbb{R}^n$, $j = 1, \dots, m$, changing the order of integration we obtain the following reformulation of Theorem 2.2.3.

Theorem 2.2.4 *Let $f_1, \dots, f_m : \mathbb{R}^k \to \mathbb{R}^+$ be measurable functions vanishing at infinity, and let $u_1, \dots, u_m \in \mathbb{R}^n$. Then,*

$$\int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} \prod_{j=1}^m f_j(\langle u_j, z_1\rangle, \dots, \langle u_j, z_k\rangle)dz_k \dots dz_1$$

$$\leq \int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} \prod_{j=1}^m f_j^*(\langle u_j, z_1\rangle, \dots, \langle u_j, z_k\rangle)dz_k \dots dz_1 . \qquad \square$$

Theorem 2.2.4 plays an important role in the original proof of the Brascamp-Lieb inequality (see Section 2.4).

2.3 Volume of restricted Minkowski sums

In this Section we present an application of the Brascamp-Lieb-Luttinger inequality: it is a Brunn-Minkowski inequality for "restricted Minkowski sums". This is then applied to yield a proof of Shannon's entropy power inequality.

2.3.1 An inequality for restricted Minkowski sums

Let A and B be non-empty subsets of $\mathbb{R}^n$ with finite Lebesgue measure. Suppose that Θ is a non-empty subset of $A \times B$. The **restricted sum** (with respect to Θ) of A and B is the set

$$A +_\Theta B = \{x + y \mid (x, y) \in \Theta\}. \tag{3.1}$$

Szarek and Voiculescu proved the following.

Theorem 2.3.1 *Let $\rho \in (0, 1)$ and let $A, B \subset \mathbb{R}^n$ with*

$$|B|^{1/n} = \rho|A|^{1/n}. \tag{3.2}$$

Assume that $\Theta \subseteq A \times B \subset \mathbb{R}^{2n}$ *satisfies*

$$|\Theta| \geq (1 - c\min\{\rho\sqrt{n}, 1\})|A| \cdot |B|. \tag{3.3}$$

Then,

$$|A +_\Theta B|^{2/n} \geq |A|^{2/n} + |B|^{2/n}. \tag{3.4}$$

Remarks: (a) In the statement above, $c > 0$ is an absolute constant (independent from ρ, n, A, B and Θ).
(b) (3.4) is weaker than the Brunn-Minkowski inequality: the point is that Θ is a "large" subset of $A \times B$, but $A +_\Theta B$ is "smaller" than the Minkowski sum $A + B$.
(c) For the proof of the theorem we may clearly assume that Θ is "maximal with respect to $A +_\Theta B$". That is,

$$\Theta = \{(x, y) \in A \times B \mid x + y \in A +_\Theta B\}. \tag{3.5}$$

(d) We may choose any normalization of volumes so that $|B|^{1/n} = \rho|A|^{1/n}$. We will assume that $|A| = |D_n|$ and $|B| = |\rho D_n|$.

Example: Let $A = D_n$, $B = \rho D_n$ and

$$\Theta = \{(x, y) \in D_n \times (\rho D_n) \mid \langle x, y\rangle \leq 0\}. \tag{3.6}$$

One can easily check that

$$|\Theta| = \frac{|A| \cdot |B|}{2}$$

and

$$A +_\Theta B = \sqrt{1 + \rho^2} D_n.$$

On the other hand,

$$|A +_\Theta B|^{2/n} = (1 + \rho^2)|D_n| = |A|^{2/n} + |B|^{2/n}. \tag{3.7}$$

This shows that the exponent $2/n$ is "best possible".

The proof of Theorem 2.3.1 is based on the Brascamp-Lieb-Luttinger rearrangement inequality (Theorem 2.2.3) and a more careful examination of the above example, which is done in the following lemma.

Lemma 2.3.1 *Let* $\rho \in (0, 1)$ *and consider the set*

$$\Theta = \left\{(x, y) \in D_n \times (\rho D_n) \mid |x + y| \leq \sqrt{1 + \rho^2}\right\}. \tag{3.8}$$

Then,

$$|\Theta| \leq (1 - c\min\{\rho\sqrt{n}, 1\})|D_n| \cdot |\rho D_n|, \tag{3.9}$$

where $c > 0$ *is an absolute constant.*

Proof: We set $\tau = \min\{\rho\sqrt{n}, 1\}/2$. We will show that if $1 - \tau/n \leq |x| \leq 1$, then

$$\left|\{y \in \rho D_n \mid |x+y| > \sqrt{1+\rho^2}\}\right| \geq c_1 |\rho D_n|, \tag{3.10}$$

where $c_1 > 0$ is an absolute constant. This implies (3.9), because

$$\begin{aligned} \left|\left(D_n \times (\rho D_n)\right)\setminus\Theta\right| &\geq \int_{\{x\in D_n : 1-\tau/n \leq |x| \leq 1\}} |\{y \in \rho D_n : |x+y| > \sqrt{1+\rho^2}\}|\, dx \\ &\geq c_1 |\rho D_n| \cdot |D_n| \left(1 - (1-\tau/n)^n\right) \\ &\geq c_2 \tau |\rho D_n| \cdot |D_n|. \end{aligned}$$

For the proof of (3.10) we may assume that $x = (r, 0, \ldots, 0)$, where $1-\tau/n \leq r \leq 1$. Drawing a picture, the reader will be convinced that the set $M := \{y \in \rho D_n \mid |x+y| \leq \sqrt{1+\rho^2}\}$ consists of two parts: the first one is

$$M_1 := \{y = (y_1, \ldots, y_n) \in \rho D_n \mid y_1 \leq s\}, \quad \text{where} \quad s = \frac{1-r^2}{2r}. \tag{3.11}$$

The volume of M_1 can be estimated by

$$|M_1| = \omega_{n-1} \int_{-\rho}^{s} (\rho^2 - u^2)^{(n-1)/2} du \leq c |\rho D_n|, \tag{3.12}$$

where $0 < c < 1$ is an absolute constant. This follows by direct computation, since

$$s = (1-r) \cdot \left(1 + \frac{1-r}{2r}\right) \leq \frac{\tau}{n} \cdot \left(1 + \frac{\tau}{2r}\right) = \frac{\rho}{2\sqrt{n}}(1 + O(1/n)). \tag{3.13}$$

The second part of M is the set

$$M_2 := \{y = (y_1, \ldots, y_n) \in \sqrt{1+\rho^2} D_n \mid r + s < y_1 \leq \sqrt{1+\rho^2}\}. \tag{3.14}$$

The volume of M_2 is equal to

$$|M_2| = \omega_{n-1} \int_{s}^{t} (1 + \rho^2 - (r+u)^2)^{(n-1)/2} du, \tag{3.15}$$

where $t = \sqrt{1+\rho^2} - r$. One can easily check that

$$|M_2| \leq (\rho/\sqrt{n}) |\rho D_n|. \tag{3.16}$$

It follows that $|M| \leq c_1 |\rho D_n|$ for some absolute constant $c_1 > 0$. $\square$

Proof of Theorem 2.3.1: We may assume that $|A| = |D_n|$, $|B| = |\rho D_n|$ and

$$\Theta = \{(x, y) \in A \times B \mid x + y \in A +_\Theta B\}.$$

Let $R > 0$ be defined by the equation

$$|A +_\Theta B| = |R D_n|. \tag{3.17}$$

Applying Theorem 2.2.3 we have

$$\begin{aligned}
|\Theta| &= |\{(x,y) \in A \times B \mid x+y \in A +_\Theta B\}| \\
&= \int\int \chi_A(x)\chi_B(y)\chi_{A+_\Theta B}(x+y)dxdy \\
&\leq \int\int \chi_A^*(x)\chi_B^*(y)\chi_{A+_\Theta B}^*(x+y)dxdy \\
&= \int\int \chi_{D_n}(x)\chi_{\rho D_n}(y)\chi_{RD_n}(x+y)dxdy \\
&= |\{(x,y) \in D_n \times (\rho D_n) \mid x+y \in RD_n\}|.
\end{aligned}$$

Let $\Theta_1 = \{(x,y) \in D_n \times (\rho D_n) \mid |x+y| \leq \sqrt{1+\rho^2}\}$. Since

$$\begin{aligned}
|\Theta_1| &\leq (1 - c\min\{\rho\sqrt{n}, 1\})|D_n| \cdot |\rho D_n| \\
&\leq |\Theta| \\
&\leq |\{(x,y) \in D_n \times (\rho D_n) \mid x+y \in RD_n\}|,
\end{aligned}$$

we get

$$R \geq \sqrt{1+\rho^2}. \tag{3.18}$$

It follows that

$$|A +_\Theta B|^{2/n} = R^2|D_n| \geq (1+\rho^2)|D_n| = |A|^{2/n} + |B|^{2/n},$$

and the proof is complete. □

Theorem 2.3.1 may be also stated in the following form.

Theorem 2.3.2 *There exist $c, C > 0$ with the following property: if $0 < \delta < c$ and A, B are subsets of $\mathbb{R}^n$ with finite Lebesgue measure, then, for any $\Theta \subset A \times B$ with*

$$|\Theta| \geq (1-\delta)|A| \cdot |B| \tag{3.19}$$

one has

$$|A +_\Theta B|^{2/n} \geq \left(1 - \frac{C\delta}{n}\right)\left(|A|^{2/n} + |B|^{2/n}\right). \tag{3.20}$$

Proof: We may normalize volumes so that

$$|A| = 1 \geq |B| = \rho^n$$

Let $c > 0$ be the constant in Theorem 2.3.1 (one may clearly assume that $c < 1/2$). We distinguish two cases.

(a) If $\rho\sqrt{n} \geq \delta/c$, then $1 - \delta \geq 1 - c\min\{\rho\sqrt{n}, 1\}$. Hence, Theorem 2.3.1 applies and

$$|A +_\Theta B|^{2/n} \geq |A|^{2/n} + |B|^{2/n}, \tag{3.21}$$

so we have the assertion of the Theorem in a stronger form.
(b) Suppose that $\rho\sqrt{n} < \delta/c$. We observe that

$$\begin{aligned}(1-\delta)|A|\cdot|B| &\leq |\Theta| \\ &= \int_B |\{x \in A \mid x+y \in A +_\Theta B\}|\, dx \\ &\leq |B|\cdot|A +_\Theta B|.\end{aligned}$$

This implies

$$|A +_\Theta B| \geq (1-\delta) \geq \left(1 - \frac{C\delta}{n}\right)(1+\rho^2) \tag{3.22}$$

if $C > 0$ is chosen large enough (independent from n). This is exactly the inequality

$$|A +_\Theta B|^{2/n} \geq \left(1 - \frac{C\delta}{n}\right)\left(|A|^{2/n} + |B|^{2/n}\right) \tag{3.23}$$

and the proof is complete. □

An alternative way of stating the result is the following.

Corollary 2.3.1 *If* $0 < \delta < \delta_0$ *and* $\Theta \subset A \times B$ *satisfies*

$$|\Theta| \geq (1-\delta)^n |A|\cdot|B|, \tag{3.24}$$

then,

$$|A +_\Theta B|^{2/n} \geq (1-\varepsilon)\left(|A|^{2/n} + |B|^{2/n}\right) \tag{3.25}$$

for some $0 < \varepsilon \leq c\sqrt{\delta}$. □

2.3.2 Shannon's inequality

Let X be an $\mathbb{R}^n$ valued random variable whose distribution is absolutely continuous with respect to the Lebesgue measure λ_n, and let f be its density. The **entropy** of X is the quantity

$$h(X) = -\int_{\mathbb{R}^n} f \log f \, d\lambda_n. \tag{3.26}$$

Shannon's **entropy power inequality** states that if X, Y are two such random variables and if X, Y are independent, then

$$\exp(2h(X)/n) + \exp(2h(Y)/n) \leq \exp(2h(X+Y)/n). \tag{3.27}$$

Shannon's inequality is one of the most important results of Information Theory. In this subsection we give a proof of this inequality, which is due to Szarek and Voiculescu. The proof is an application of the restricted Minkowski sums estimate in the form of Corollary 2.3.1.

We consider only the real valued case. Let $X_1, X_2, \ldots, X_N, \ldots$ be a sequence of independent copies of X on some probability space (Ω, P). For every $N \in \mathbb{N}$, the joint density F_N of $X_1, X_2, \ldots, X_N$ with respect to Lebesgue measure is the function

$$F(x_1, x_2, \ldots, x_N) = f(x_1) f(x_2) \ldots f(x_N). \tag{3.28}$$

Proposition 2.3.1 *There exist $\alpha_k, \beta_k > 0$ with $\alpha_k \to 0$ and $\beta_k \to 0$, such that*

$$P((X_1, X_2, \ldots, X_N) \in V_N(X)) > 1 - \beta_N, \tag{3.29}$$

where

$$V_N(X) = \{x \in \mathbb{R}^N \mid \exp(N(-h(X) - \alpha_N)) \leq F_N(x) \leq \exp(N(-h(X) + \alpha_N)\}. \tag{3.30}$$

Proof: Consider the function

$$\frac{\log F_N}{N}(\overline{x}) := \frac{1}{N} \sum_{j=1}^{N} \log f(x_j) \tag{3.31}$$

as a random variable on $\mathbb{R}^\infty$ equipped with the product measure $P := \otimes_{j=1}^{\infty} (f d\lambda_1)$, Then,

$$\frac{\log F_N}{N} = \frac{1}{N} \sum_{j=1}^{N} \log f(X_j).$$

By the law of large numbers, $\log F_N / N$ converges in probability to

$$\mathbb{E}(\log f(X)) = \int f \log f \, d\lambda_1 = -h(X). \tag{3.32}$$

Therefore, there exists a sequence (α_N) of positive reals with $\lim_N \alpha_N = 0$, such that

$$P\left(-h(X) - \alpha_N \leq \frac{1}{N} \sum_{j=1}^{N} \log f(X_j) \leq -h(X) + \alpha_N\right) \to 1$$

as $N \to \infty$. This is equivalent to the statement of the Proposition. $\square$

Corollary 2.3.2 *In the notation of Proposition 2.3.1,*

$$(1 - \beta_N) e^{N(h(X) - \alpha_N)} \leq |V_N(X)| \leq e^{N(h(X) + \alpha_N)}. \tag{3.33}$$

Proof: We just observe that

$$P((X_1, X_2, \ldots, X_N) \in V_N(X)) = \int_{V_N(X)} F_N(x)\, dx > 1 - \beta_N, \tag{3.34}$$

and use the definition of $V_N(X)$. □

Letting $N \to \infty$ we get

$$\lim_{N\to\infty} \frac{\log |V_N(X)|}{N} = h(X) \tag{3.35}$$

and

$$\lim_{N\to\infty} |V_N(X)|^{2/N} = \exp(2h(X)). \tag{3.36}$$

Theorem 2.3.3 *Let X, Y be independent random variables with densities f, g with respect to λ_1. Then,*

$$\exp(2h(X)) + \exp(2h(Y)) \leq \exp(2h(X+Y)).$$

Proof: Let (X_k) and (Y_k) be sequences of jointly independent copies of X and Y respectively. We define $A_N = V_N(X)$, $B_N = V_N(Y)$ and

$$\Theta_N = \{(x,y) \in A_N \times B_N \mid x + y \in V_N(X+Y)\}. \tag{3.37}$$

We may assume that Proposition 2.3.1 holds true for A_N, B_N and $C_N := V_N(X+Y)$ simultaneously (i.e. with the same sequences (α_k), (β_k)).

If $\mathbf{X}_N = (X_1, \ldots, X_N)$ and $\mathbf{Y}_N = (Y_1, \ldots, Y_N)$, then

$$\begin{aligned}
1 - \beta_N &\leq P(\mathbf{X}_N + \mathbf{Y}_N \in C_N) \\
&\leq P(\mathbf{X}_N \in A_N\ ,\ \mathbf{Y}_N \in B_N, \mathbf{X}_N + \mathbf{Y}_N \in C_N) \\
&\quad + P((\mathbf{X}_N, \mathbf{Y}_N) \notin A_N \times B_N) \\
&= P((\mathbf{X}_N, \mathbf{Y}_N) \in \Theta_N) + 1 - P(\mathbf{X}_N \in A_N) \cdot P(\mathbf{Y}_N \in B_N) \\
&\leq P((\mathbf{X}_N, \mathbf{Y}_N) \in \Theta_N) + 1 - (1 - \beta_N)^2.
\end{aligned}$$

Therefore, if F_N, G_N are the densities of $\mathbf{X}_N, \mathbf{Y}_N$ with respect to λ_N, we have

$$\int_{\Theta_N} F_N G_N d\lambda_N = P((\mathbf{X}_N, \mathbf{Y}_N) \in \Theta_N) \geq 1 - 3\beta_N + \beta_N^2. \tag{3.38}$$

By the definition of A_N, B_N we have

$$\int_{\Theta_N} F_N G_N d\lambda_N \leq |\Theta_N| \cdot e^{N(-h(X)+\alpha_N)} e^{N(-h(Y)+\alpha_N)}, \tag{3.39}$$

while Corollary 2.3.2 shows that

$$|A_N| \cdot |B_N| \leq e^{N(h(X)+\alpha_N)} e^{N(h(Y)+\alpha_N)}. \tag{3.40}$$

From (3.38), (3.39) and (3.40) we see that

$$\frac{|\Theta_N|}{|A_N|\cdot|B_N|} \geq (1-3\beta_N+\beta_N^2)e^{-4N\alpha_N}, \tag{3.41}$$

which shows that

$$\lim_{N\to\infty}\left(\frac{|\Theta_N|}{|A_N|\cdot|B_N|}\right)^{1/N} = 1. \tag{3.42}$$

Now, Corollary 2.3.1 shows that

$$|C_N|^{2/N} = |A_N +_{\Theta_N} B_N|^{2/N} \geq (1-\varepsilon_N)\left(|A_N|^{2/N}+|B_N|^{2/N}\right) \tag{3.43}$$

where $\varepsilon_N \to 0^+$ as $N\to\infty$. Taking into account (3.36) we conclude the proof. □

Remark: The same proof gives the inequality

$$\exp(2h(X)/n)+\exp(2h(Y)/n) \leq \exp(2h(X+Y)/n$$

if X and Y are $\mathbb{R}^n$ valued. Then, $V_N(X)$ and $V_N(Y)$ are subsets of $\mathbb{R}^{Nn}$, we have $|V_N(X)|^{2/(Nn)} \to \exp(2h(X)/n)$, etc.

2.4 Brascamp-Lieb inequality

Let $m\geq n$, $p_1,\ldots,p_m\geq 1$ with $\frac{1}{p_1}+\cdots+\frac{1}{p_m}=n$, and $u_1,\ldots,u_m\in\mathbb{R}^n$. Consider the multilinear operator $\Phi: L^{p_1}(\mathbb{R})\times\cdots\times L^{p_m}(\mathbb{R})\to\mathbb{R}$ defined by

$$\Phi(f_1,\ldots,f_m) = \int_{\mathbb{R}^n}\prod_{j=1}^m f_j(\langle u_j,x\rangle)\,dx. \tag{4.1}$$

Brascamp and Lieb proved that the norm of Φ is the supremum of

$$\frac{\Phi(g_1,\ldots,g_m)}{\prod_{j=1}^m\|g_j\|_{p_j}} \tag{4.2}$$

over all centered Gaussian functions $g_1,\ldots,g_m$. The proof of this fact is based on the Brascamp-Lieb-Luttinger rearrangement inequality (Theorem 2.2.4).

Theorem 2.4.1 *Let $m\geq n$, and $p_1,\ldots,p_m\geq 1$ with $\frac{1}{p_1}+\cdots+\frac{1}{p_m}=n$. Given $u_1,\ldots,u_m\in\mathbb{R}^n$, consider the operator Φ (depending on the u_j's). Then,*

$$\frac{\Phi(f_1,\ldots,f_m)}{\prod_{j=1}^m\|f_j\|_{p_j}} \leq D := \sup\left\{\frac{\Phi(g_1,\ldots,g_m)}{\prod_{j=1}^m\|g_j\|_{p_j}} : g_j(t)=e^{-\lambda_j t^2}, \lambda_j>0\right\} \tag{4.3}$$

for all $f_j\in L^{p_j}(\mathbb{R})$.

Remarks: (a) We may clearly assume that the f_j's are non-negative.

(b) We saw that the symmetric decreasing rearrangement preserves p-norms and increases the left hand side in Theorem 2.4.1. Therefore, we may assume that each f_j is even and decreasing on $\mathbb{R}^+$. Such a function can be approximated by functions of the form

$$\sum_{l=1}^{k} a_l \chi_{[-\gamma_l, \gamma_l]}$$

where $a_l > 0$ and γ_l are decreasing. Hence, it suffices to prove the theorem for functions of this type.

Lemma 2.4.1 *Let $\psi_1, \ldots, \psi_s$ be non-negative functions in $L^p(\mathbb{R}^k)$, $p \geq 1$. Then,*

$$\left\| \sum_{j=1}^{s} \psi_j \right\|_p \geq s^{-1/q} \sum_{j=1}^{s} \|\psi_j\|_p, \tag{4.4}$$

where $\frac{1}{p} + \frac{1}{q} = 1$.

Proof: Since each ψ_j is non-negative and $p \geq 1$, for every $x \in \mathbb{R}^k$ we have

$$\sum_{j=1}^{s} \psi_j(x)^p \leq \left(\sum_{j=1}^{s} \psi_j(x) \right)^p .$$

It follows that

$$\begin{aligned}
\left(\sum_{j=1}^{s} \|\psi_j\|_p \right)^p &\leq s^{p/q} \sum_{j=1}^{s} \|\psi_j\|_p^p = s^{p/q} \int_{\mathbb{R}^k} \sum_{j=1}^{s} \psi_j^p(x) dx \\
&\leq s^{p/q} \int_{\mathbb{R}^k} \left(\sum_{j=1}^{s} \psi_j(x) \right)^p dx \\
&= s^{p/q} \left\| \sum_{j=1}^{s} \psi_j \right\|_p^p . \qquad \square
\end{aligned}$$

Lemma 2.4.2 *Let $p \geq 1$, and let η_a be the characteristic function of the ball $\{x \in \mathbb{R}^k : |x| \leq a\}$. We define*

$$\phi_a(x) = \exp\left(\left(1 - \frac{|x|^2}{a^2} \right) \frac{k}{2p} \right) .$$

Then, $\eta_a(x) \leq \phi_a(x)$ for every $x \in \mathbb{R}^k$ and

$$\|\phi_a\|_p \leq \|\eta_a\|_p (3\sqrt{k})^{1/p}. \tag{4.5}$$

Proof: The inequality $\eta_a(x) \leq \phi_a(x)$ is obvious since $\phi_a(x) \geq 1$ when $|x| \leq a$. We have $\|\eta_a\|_p^p = \omega_k a^k$ and

$$\begin{aligned}
\|\phi_a\|_p^p &= k\omega_k \int_0^\infty t^{k-1} e^{(1-t^2/a^2)\frac{k}{2}} dt \\
&= k\omega_k e^{k/2} \int_0^\infty t^{k-1} e^{-\frac{t^2 k}{2a^2}} dt \\
&= \omega_k e^{k/2} a^k \left(\frac{2}{k}\right)^{k/2} \Gamma\left(\frac{k}{2}+1\right) \\
&= \|\eta_a\|_p^p e^{k/2} \left(\frac{2}{k}\right)^{k/2} \Gamma\left(\frac{k}{2}+1\right).
\end{aligned}$$

Using the precise form of Stirling's formula we obtain

$$\left(\frac{\|\phi_a\|_p}{\|\eta_a\|_p}\right)^p \leq 3\sqrt{k}. \qquad \square$$

Remark: Observe that ϕ_a is a product of centered Gaussian functions:

$$\phi_a(x_1,\ldots,x_k) = e^{k/(2p)} \prod_{i=1}^k \exp\big(-\frac{kx_i^2}{2a^2p}\big).$$

Proof of the Brascamp-Lieb inequality: We may assume that each f_j is a function of the form

$$f_j(t) = \sum_{l=1}^A a_{lj}\chi_{[-\gamma_{lj},\gamma_{lj}]}(t)$$

where $a_{lj} \geq 0$, $\gamma_{lj} \geq \gamma_{l+1,j}$ and A is a positive integer (the same for all f_j).
For every $j \leq m$ we consider the function $G_j : \mathbb{R}^k \to \mathbb{R}^+$ defined by

$$G_j(t_1,\ldots,t_k) = \prod_{i=1}^k f_j(t_i),$$

and the symmetric decreasing rearrangement F_j of G_j. Theorem 2.2.4 and Fubini's theorem show that

$$\begin{aligned}
\left(\int_{\mathbb{R}^n} \prod_{j=1}^m f_j(\langle u_j, x\rangle)dx\right)^k &= \int_{\mathbb{R}^{nk}} \prod_{j=1}^m G_j(\langle u_j, x_1\rangle,\ldots,\langle u_j, x_k\rangle)dx_k\ldots dx_1 \\
&\leq \int_{\mathbb{R}^{nk}} \prod_{j=1}^m F_j(\langle u_j, x_1\rangle,\ldots,\langle u_j, x_k\rangle)dx_k\ldots dx_1.
\end{aligned}$$

Since f_j takes only A values, G_j, and so F_j as well, takes at most $(k+1)^A$ values (if $\lambda_1,\ldots,\lambda_A$ are the values of f_j, then the values of G_j are $\lambda_1^{r_1}\ldots\lambda_A^{r_A}$ where $0 \leq r_l \leq k$

and $\sum r_l = k$: so, we estimate by the number of distinct A-tuples of elements of $\{0, 1, \ldots, k\}$). It follows that each F_j is a function of the form

$$F_j = \sum_{l=1}^{(k+1)^A} H_{lj}\eta_{a_{lj}}, \tag{4.6}$$

where $H_{lj} \geq 0$ and $\{a_{lj}\}$ is decreasing. From Lemma 2.4.1,

$$\|f_j\|_{p_j}^k = \|F_j\|_{p_j} = \Big\| \sum_{l=1}^{(k+1)^A} H_{lj}\eta_{a_{lj}} \Big\|_{p_j} \geq (k+1)^{-A/q_j} \sum_{l=1}^{(k+1)^A} H_{lj}\|\eta_{a_{lj}}\|_{p_j}, \tag{4.7}$$

where q_j is the conjugate exponent of p_j. We write

$$\begin{aligned} \left(\frac{\Phi(f_1, \ldots, f_m)}{\prod_{j=1}^m \|f_j\|_{p_j}}\right)^k &= \frac{\int_{\mathbb{R}^{nk}} \prod_{j=1}^m G_j(\langle u_j, x_1\rangle, \ldots, \langle u_j, x_k\rangle)dx_k \ldots dx_1}{\prod_{j=1}^m \|G_j\|_{p_j}} \\ &\leq \frac{\int_{\mathbb{R}^{nk}} \prod_{j=1}^m F_j(\langle u_j, x_1\rangle, \ldots, \langle u_j, x_k\rangle)dx_k \ldots dx_1}{\prod_{j=1}^m \|F_j\|_{p_j}}. \end{aligned}$$

From (4.6) and (4.7) this is bounded by

$$J := \frac{\sum_{l_1} \ldots \sum_{l_m} H_{l_1,1} \ldots H_{l_m,m} \int_{\mathbb{R}^{nk}} \prod_{j=1}^m \eta_{a_{l_j,j}}(\langle u_j, x_1\rangle, \ldots, \langle u_j, x_k\rangle)dx_k \ldots dx_1}{(k+1)^{-A\sum \frac{1}{q_j}} \sum_{l_1} \ldots \sum_{l_m} H_{l_1,1} \ldots H_{l_m,m} \prod_{j=1}^m \|\eta_{a_{l_j,j}}\|_{p_j}}.$$

Let us consider an m-tuple $(\eta_{b_1}, \ldots, \eta_{b_m})$ and the functions $(\phi_{b_1}, \ldots, \phi_{b_m})$ as in Lemma 2.4.2. We have $\eta_{b_j} \leq \phi_{b_j}$, and hence

$$\int_{\mathbb{R}^{nk}} \prod_{j=1}^m \eta_{b_j}(\langle u_j, x_1\rangle, \ldots, \langle u_j, x_k\rangle) \leq \int_{\mathbb{R}^{nk}} \prod_{j=1}^m \phi_{b_j}(\langle u_j, x_1\rangle, \ldots, \langle u_j, x_k\rangle). \tag{4.8}$$

Also,

$$\prod_{j=1}^m \|\phi_{b_j}\|_{p_j} \leq (3\sqrt{k})^{\sum \frac{1}{p_j}} \prod_{j=1}^m \|\eta_{b_j}\|_{p_j} = (3\sqrt{k})^n \prod_{j=1}^m \|\eta_{b_j}\|_{p_j}. \tag{4.9}$$

Each ϕ_{b_j} is a product of k identical one dimensional centered Gaussians $\phi_{b_j}^{(l)}$ evaluated at the points $\langle u_j, x_l\rangle$ respectively. Therefore, the right hand side of (4.8) is the k-th power of the integrals on $\mathbb{R}^n$ of the product of m centered Gaussians evaluated at $\langle u_j, x\rangle$, $j = 1, \ldots, m$. Since

$$\|\phi_{b_j}\|_{p_j} = \|\prod_{l=1}^k \phi_{b_j}^{(l)}\|_{p_j} = \prod_{l=1}^k \|\phi_{b_j}^{(l)}\|_{p_j}, \tag{4.10}$$

it follows that

$$\frac{\int_{\mathbb{R}^{nk}}\prod_{j=1}^m \phi_{b_j}(\langle u_j,x_1\rangle,\dots,\langle u_j,x_k\rangle)}{\prod_{j=1}^m \|\phi_{b_j}\|_{p_j}} \leq D^k. \tag{4.11}$$

Combining the above, we obtain

$$\begin{aligned}\left(\frac{\Phi(f_1,\dots,f_m)}{\prod_{j=1}^m \|f_j\|_{p_j}}\right)^k &\leq J \leq (k+1)^{A\sum_{j=1}^m \frac{1}{q_j}}(3\sqrt{k})^n D^k \\ &= (k+1)^{(m-n)A}(3\sqrt{k})^n D^k.\end{aligned}$$

Then,

$$\frac{\Phi(f_1,\dots,f_m)}{\prod_{j=1}^m \|f_j\|_{p_j}} \leq (k+1)^{\frac{(m-n)A}{k}}(3\sqrt{k})^{\frac{n}{k}} D, \tag{4.12}$$

and letting $k\to\infty$, we conclude the proof. □

2.5 Reverse Brascamp-Lieb inequality

Recall the statement of the Brascamp-Lieb inequality for non-negative functions: If $m\geq n$, $u_1,\dots,u_m\in\mathbb{R}^n$ and $p_1,\dots,p_m\geq 1$ with $\frac{1}{p_1}+\cdots+\frac{1}{p_m}=n$, then

$$\int_{\mathbb{R}^n}\prod_{j=1}^m f_j(\langle x,u_j\rangle)dx \leq D\cdot\prod_{j=1}^m \|f_j\|_{p_j} \tag{5.1}$$

for all non-negative $f_j\in L^{p_j}(\mathbb{R})$, where

$$D=\sup\left\{\frac{\int_{\mathbb{R}^n}\prod_{j=1}^m g_j(\langle x,u_j\rangle)dx}{\prod_{j=1}^m \|g_j\|_{p_j}} : g_j(t)=e^{-\lambda_j t^2}, \lambda_j>0\right\}. \tag{5.2}$$

If we set $c_j=1/p_j$ and replace f_j by $f_j^{c_j}$ we obtain the following:

Theorem 2.5.1 *If $m\geq n$, $u_1,\dots,u_m\in\mathbb{R}^n$ and $c_1,\dots,c_m>0$ with $c_1+\cdots+c_m=n$, then*

$$\int_{\mathbb{R}^n}\prod_{j=1}^m f_j^{c_j}(\langle x,u_j\rangle)dx \leq D\cdot\prod_{j=1}^m\left(\int_{\mathbb{R}} f_j\right)^{c_j} \tag{5.3}$$

for all integrable functions $f_j:\mathbb{R}\to\mathbb{R}^+$. □

The value of the constant D is described by the following proposition.

Proposition 2.5.1 *Let $m \geq n$, $c_1, \ldots, c_m > 0$ with $c_1 + \cdots + c_m = n$, and $u_1, \ldots, u_m \in \mathbb{R}^n$. Then, $D = 1/\sqrt{F}$ where*

$$F = \inf\left\{ \frac{\det\left(\sum_{j=1}^m c_j \lambda_j u_j \otimes u_j\right)}{\prod_{j=1}^m \lambda_j^{c_j}} \mid \lambda_j > 0 \right\}. \tag{5.4}$$

Proof: Recall that $(u \otimes u)(x) = \langle x, u \rangle u$. Let $g_j(t) = \exp(-\lambda_j t^2)$, $j = 1, \ldots, m$, where λ_j are positive reals. Then,

$$\begin{aligned}
\int_{\mathbb{R}^n} \prod_{j=1}^m g_j^{c_j}(\langle x, u_j \rangle) dx &= \int_{\mathbb{R}^n} \exp\left(-\sum_{j=1}^m c_j \lambda_j \langle x, u_j \rangle^2 \right) dx \\
&= \int_{\mathbb{R}^n} \exp\left(-\big\langle \big(\sum_{j=1}^m c_j \lambda_j u_j \otimes u_j \big)(x), x \big\rangle \right) dx \\
&= \frac{\pi^{n/2}}{\sqrt{\det\left(\sum_{j=1}^m c_j \lambda_j u_j \otimes u_j\right)}}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\prod_{j=1}^m \left(\int_{\mathbb{R}} g_j \right)^{c_j} &= \prod_{j=1}^m \left(\int_{\mathbb{R}} \exp(-\lambda_j t^2) dt \right)^{c_j} \\
&= \prod_{j=1}^m \left(\frac{\sqrt{\pi}}{\sqrt{\lambda_j}} \right)^{c_j} \\
&= \frac{\pi^{n/2}}{\sqrt{\prod_{j=1}^m \lambda_j^{c_j}}}
\end{aligned}$$

since $c_1 + \cdots + c_m = n$. It follows that

$$\begin{aligned}
\frac{1}{D^2} &= \inf\left\{ \left(\frac{\prod_{j=1}^m \left(\int_{\mathbb{R}} g_j\right)^{c_j}}{\int_{\mathbb{R}^n} \prod_{j=1}^m g_j^{c_j}(\langle x, u_j \rangle) dx} \right)^2 \mid g_j(t) = e^{-\lambda_j t^2},\ \lambda_j > 0 \right\} \\
&= \inf\left\{ \frac{\det\left(\sum_{j=1}^m c_j \lambda_j u_j \otimes u_j\right)}{\prod_{j=1}^m \lambda_j^{c_j}} \mid \lambda_j > 0 \right\}. \qquad \square
\end{aligned}$$

In the sequel we do not assume the Brascamp-Lieb inequality. The computation in Proposition 2.5.1 shows the following.

Lemma 2.5.1 *Let $m \geq n$, $c_1, \ldots, c_m > 0$ with $c_1 + \cdots + c_m = n$, and $u_1, \ldots, u_m \in \mathbb{R}^n$. If $f_1, \ldots, f_m : \mathbb{R} \to \mathbb{R}^+$ are measurable functions, let*

$$I(f_1, \ldots, f_m) = \int_{\mathbb{R}^n} \prod_{j=1}^m f_j^{c_j}(\langle x, u_j \rangle) dx. \tag{5.5}$$

Then,

$$\sup\left\{I(f_1,\ldots,f_m)\mid \int_{\mathbb{R}} f_j=1\ ,\ j=1,\ldots,m\right\}\geq \frac{1}{\sqrt{F}}. \tag{5.6}$$

Proof: Obvious, by homogenuity. The Brascamp-Lieb inequality in the form of Theorem 2.5.1 asserts that there is equality in (5.6). Here, we just use centered Gaussians in order to check the easy part. □

Barthe proved the following reverse form of Theorem 2.5.1 (this was inspired by applications in Convex Geometric Analysis and was conjectured by Ball; we will discuss the applications in detail in subsequent chapters of these notes).

Theorem 2.5.2 *Let $m\geq n$, $c_1,\ldots,c_m>0$ with $c_1+\cdots+c_m=n$, and $u_1,\ldots,u_m\in\mathbb{R}^n$. If $h_1,\ldots,h_m:\mathbb{R}\to\mathbb{R}^+$ are measurable functions, we set*

$$K(h_1,\ldots,h_m)=\int_{\mathbb{R}^n}\sup\left\{\prod_{j=1}^m h_j^{c_j}(\theta_j)\mid \theta_j\in\mathbb{R}\ ,\ x=\sum_{j=1}^m\theta_jc_ju_j\right\}dx. \tag{5.7}$$

Then,

$$\inf\left\{K(h_1,\ldots,h_m)\mid \int_{\mathbb{R}} h_j=1\ ,\ j=1,\ldots,m\right\}=\sqrt{F}. \tag{5.8}$$

Testing centered Gaussian functions as in Proposition 2.5.1, we get the easy part of this **reverse Brascamp-Lieb inequality**.

Proposition 2.5.2 *With the notation of Theorem 2.5.2 we have*

$$\inf\left\{K(h_1,\ldots,h_m)\mid \int_{\mathbb{R}} h_j=1\ ,\ j=1,\ldots,m\right\}\leq\sqrt{F}. \tag{5.9}$$

Proof: Let $\lambda_j>0$, $j=1,\ldots,m$ and consider the functions $h_j(t)=\exp(-t^2/\lambda_j)$. Then, the function

$$m(x):=\sup\left\{\prod_{j=1}^m h_j^{c_j}(\theta_j)\mid x=\sum_{j=1}^m\theta_jc_ju_j\right\}$$

is given by

$$m(x)=\exp\left(-\inf\left\{\sum_{j=1}^m\frac{c_j}{\lambda_j}\theta_j^2\mid x=\sum_{j=1}^m\theta_jc_ju_j\right\}\right).$$

Define

$$\|x\|^2=\sum_{j=1}^m c_j\lambda_j\langle x,u_j\rangle^2=\langle Ax,x\rangle \tag{5.10}$$

where A is the symmetric positive-definite operator $A := \sum_{j=1}^m c_j\lambda_j u_j \otimes u_j$. It is not hard to check that the dual norm is exactly

$$\|y\|_*^2 = \inf\left\{\sum_{j=1}^m \frac{c_j}{\lambda_j}\theta_j^2 \mid y = \sum_{j=1}^m \theta_j c_j u_j\right\}. \tag{5.11}$$

Therefore,

$$\|y\|_*^2 = \langle By, y\rangle, \tag{5.12}$$

where $B = A^{-1}$. It follows that

$$\int_{\mathbb{R}^n} m(x)dx = \frac{\pi^{n/2}}{\sqrt{\det B}} = \pi^{n/2}\sqrt{\det A}. \tag{5.13}$$

On the other hand,

$$\prod_{j=1}^m \left(\int_{\mathbb{R}} \exp(-t^2/\lambda_j)dt\right)^{c_j} = \pi^{n/2}\prod_{j=1}^m \lambda_j^{c_j/2}. \tag{5.14}$$

This shows that

$$\begin{aligned}\inf\left\{K^2(h_1,\ldots,h_m) \mid \int_{\mathbb{R}} h_j = 1\right\} &\leq \inf\left\{\frac{\det\left(\sum_{j=1}^m c_j\lambda_j u_j\otimes u_j\right)}{\prod_{j=1}^m \lambda_j^{c_j}} \mid \lambda_j > 0\right\} \\ &= F\end{aligned}$$

and the proof is complete. $\square$

The main step in Barthe's argument is the following proposition.

Proposition 2.5.3 *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,

$$F\cdot I(f_1,\ldots,f_m) \leq K(h_1,\ldots,h_m). \tag{5.15}$$

Proof: We may assume that f_j, h_j are continuous and strictly positive. We may also assume that $0 < F < +\infty$ (F is not degenerated). We use the transportation of measure idea that was used for the proof of the Prékopa-Leindler inequality: For every $j = 1,\ldots,m$ we define $T_j : \mathbb{R}\to\mathbb{R}$ by the equation

$$\int_{-\infty}^{T_j(t)} h_j(s)ds = \int_{-\infty}^{t} f_j(s)ds. \tag{5.16}$$

Then, each T_j is strictly increasing, 1-1 and onto, and

$$T_j'(t)h_j(T_j(t)) = f_j(t), \quad t \in \mathbb{R}. \tag{5.17}$$

We now define $W : \mathbb{R}^n \to \mathbb{R}^n$ by

$$W(y) = \sum_{j=1}^{m} c_j T_j(\langle y, u_j\rangle)u_j. \tag{5.18}$$

A simple computation shows that

$$J(W)(y) = \sum_{j=1}^{m} c_j T_j'(\langle y, u_j\rangle)u_j \otimes u_j. \tag{5.19}$$

This shows that

$$\langle [J(W)(y)](v), v\rangle > 0 \quad \text{if} \quad v \neq 0$$

and hence, W is injective. Consider the function

$$m(x) = \sup\left\{ \prod_{j=1}^{m} h_j^{c_j}(\theta_j) \mid x = \sum_{j=1}^{m} \theta_j c_j u_j \right\}.$$

Then, (5.18) shows that

$$m(W(y)) \geq \prod_{j=1}^{m} h_j^{c_j}(T_j(\langle y, u_j\rangle)) \tag{5.20}$$

for every $y \in \mathbb{R}^n$. It follows that

$$\begin{aligned}
\int_{\mathbb{R}^n} m(x)dx &\geq \int_{W(\mathbb{R}^n)} m(x)dx \\
&= \int_{\mathbb{R}^n} m(W(y)) \cdot |J(W)(y)|\, dy \\
&\geq \int_{\mathbb{R}^n} \prod_{j=1}^{m} h_j^{c_j}(T_j(\langle y, u_j\rangle)) \det\left(\sum_{j=1}^{m} c_j T_j'(\langle y, u_j\rangle)u_j \otimes u_j\right) dy.
\end{aligned}$$

By the definition of F we have

$$\det\left(\sum_{j=1}^{m} c_j T_j'(\langle y, u_j\rangle)u_j \otimes u_j\right) \geq F \cdot \prod_{j=1}^{m} \left(T_j'(\langle y, u_j\rangle)\right)^{c_j}. \tag{5.21}$$

Therefore, taking (5.17) into account we have

$$\begin{aligned}
\int_{\mathbb{R}^n} m(x)dx &\geq F \cdot \int_{\mathbb{R}^n} \prod_{j=1}^{m} h_j^{c_j}(T_j(\langle y, u_j\rangle)) \cdot \prod_{j=1}^{m} \left(T_j'(\langle y, u_j\rangle)\right)^{c_j} dy \\
&= F \cdot \int_{\mathbb{R}^n} \prod_{j=1}^{m} f_j^{c_j}(\langle y, u_j\rangle)dy \\
&= F \cdot I(f_1, \ldots, f_m).
\end{aligned}$$

In other words, $F\cdot I(f_1,\dots,f_m)\le K(h_1,\dots,h_m)$. □

We can now prove simultaneously the Brascamp-Lieb inequality (Theorem 2.5.1) and the reverse Brascamp-Lieb inequality (Theorem 2.5.2).

Theorem 2.5.3 *Let $m\ge n$, $c_1,\dots,c_m>0$ with $c_1+\cdots+c_m=n$, and $u_1,\dots,u_m\in\mathbb{R}^n$. If $f_1,\dots,f_m:\mathbb{R}\to\mathbb{R}^+$ and $h_1,\dots,h_m:\mathbb{R}\to\mathbb{R}^+$ are integrable functions with*

$$\int_{\mathbb{R}} f_j(t)dt=\int_{\mathbb{R}} h_j(t)dt=1,\quad j=1,\dots,m,$$

then,

$$I(f_1,\dots,f_m)\le\frac{1}{\sqrt{F}} \tag{5.22}$$

and

$$K(h_1,\dots,h_m)\ge F. \tag{5.23}$$

Proof: Combining Lemma 2.5.1, Proposition 2.5.2 and Theorem 2.5.3, we have

$$\begin{aligned}\frac{1}{\sqrt{F}} &\le \sup\left\{I(f_1,\dots,f_m)\mid\int_{\mathbb{R}} f_j=1\right\}\\ &\le \frac{1}{F}\cdot\inf\left\{K(h_1,\dots,h_m)\mid\int_{\mathbb{R}} h_j=1\right\}\\ &\le \frac{1}{\sqrt{F}}.\end{aligned}$$

We must have equality everywhere, and this proves the theorem. □

We close this Section with a definition and a remark which will be very important for the applications to Convex Geometry.

Definition: Let $m\ge n$. We say that the vectors $u_1,\dots,u_m\in S^{n-1}$ are in **isotropic position with weights** $c_1,\dots,c_m$, if the identity operator I is decomposed in the form

$$I=\sum_{j=1}^m c_ju_j\otimes u_j. \tag{5.24}$$

Taking traces in both sides we see that

$$c_1+\cdots+c_m=n \tag{5.25}$$

is a necessary condition for (5.24). It is easy to check that the isotropic condition (5.24) is equivalent to the following: For every $x\in\mathbb{R}^n$,

$$\sum_{j=1}^m c_j\langle x,u_j\rangle^2=|x|^2. \tag{5.26}$$

Theorem 2.5.4 *Assume that the vectors $u_1, \dots, u_m \in S^{n-1}$ are in isotropic position with weights $c_1, \dots, c_m > 0$. Then,*

$$F = F(\{u_j\}, \{c_j\}) = 1, \tag{5.27}$$

where F is the constant in the Brascamp-Lieb inequality and its reverse for this set of u_j's and c_j's.

Proof: Let $\lambda_j > 0$, $j = 1, \dots, m$. For every $I \subseteq \{1, \dots, m\}$ with cardinality $|I| = n$ we define

$$\lambda_I = \prod_{i \in I} \lambda_j \quad \text{and} \quad U_I = \big(\det\big(\sqrt{c_j} u_j : j \in I\big)\big)^2. \tag{5.28}$$

By the Cauchy-Binet formula we have

$$\det\left(\sum_{j=1}^m c_j \lambda_j u_j \otimes u_j\right) = \det\left(\sum_{j=1}^m \lambda_j (\sqrt{c_j} u_j) \otimes (\sqrt{c_j} u_j)\right) = \sum_{|I|=n} \lambda_I U_I. \tag{5.29}$$

Setting $\lambda_j = 1$ in (5.29) and using the isotropic condition, we see that

$$\sum_{|I|=n} U_I = 1. \tag{5.30}$$

By the arithmetic-geometric means inequality,

$$\sum_{|I|=n} \lambda_I U_I \geq \prod_{|I|=n} \lambda_I^{U_I} = \prod_{j=1}^m \lambda_j^{\sum_{\{I : j \in I\}} U_I}. \tag{5.31}$$

Applying the Cauchy-Binet formula again, we have

$$\begin{aligned}
\sum_{\{I : j \in I\}} U_I &= \sum_{|I|=n} U_I - \sum_{\{I : j \notin I\}} U_I \\
&= 1 - \det\big(I - (\sqrt{c_j} u_j) \otimes (\sqrt{c_j} u_j)\big) \\
&= 1 - (1 - c_j |u_j|^2) \\
&= c_j.
\end{aligned}$$

Going back to (5.29) and (5.30) we see that

$$\det\left(\sum_{j=1}^m c_j \lambda_j u_j \otimes u_j\right) \geq \prod_{j=1}^m \lambda_j^{c_j}. \tag{5.32}$$

Since the λ_j's were arbitrary, we conclude that $F \geq 1$. The choice $\lambda_j = 1$ gives equality in (5.32). Therefore, $F = 1$. $\square$

2.6 Multidimensional versions

Let $S^+(\mathbb{R}^k)$ be the set of $k\times k$ symmetric, positive definite matrices. If $A \in S^+(\mathbb{R}^k)$, we write G_A for the centered Gaussian function $G_A : \mathbb{R}^k \to \mathbb{R}$ defined by

$$G_A(x) = \exp(-\langle Ax, x\rangle). \tag{6.1}$$

Finally, we write $L_1^+(\mathbb{R}^k)$ for the class of integrable non-negative functions $f : \mathbb{R}^k \to \mathbb{R}$. Our purpose in this section is to prove the multidimensional version of the Brascamp-Lieb inequality and its reverse. The setting is the following:

Let $m \geq n$. Suppose we are given real numbers $c_1, \dots, c_m > 0$ and integers $n_1, \dots, n_m$ less than or equal to n, such that

$$\sum_{j=1}^m c_j n_j = n. \tag{6.2}$$

For each $j = 1, \dots, m$, we are also given a linear map $B_j : \mathbb{R}^n \to \mathbb{R}^{n_j}$ which is onto. We also assume that

$$\bigcap_{j=1}^m \operatorname{Ker}(B_j) = \{0\}. \tag{6.3}$$

We define two operators $I, K : L_1^+(\mathbb{R}^{n_1}) \times \cdots \times L_1^+(\mathbb{R}^{n_m}) \to \mathbb{R}$ by

$$I(f_1, \dots, f_m) = \int_{\mathbb{R}^n} \prod_{j=1}^m f_j^{c_j}(B_j x)dx \tag{6.4}$$

and

$$K(h_1, \dots, h_m) = \int_{\mathbb{R}^m}^* m(x)dx, \tag{6.5}$$

where

$$m(x) = \sup\left\{ \prod_{j=1}^m h_j^{c_j}(y_j) \mid y_j \in \mathbb{R}^{n_j} \ \text{ and } \ \sum_{j=1}^m c_j B_j^* y_j = x \right\}. \tag{6.6}$$

Let E be the largest constant for which

$$K(h_1, \dots, h_m) \geq E \cdot \prod_{j=1}^m \left(\int_{\mathbb{R}^{n_j}} h_j \right)^{c_j} \tag{6.7}$$

holds true for all $h_j \in L_1^+(\mathbb{R}^{n_j})$, and let F be the smallest constant for which

$$I(f_1, \dots, f_m) \leq F \cdot \prod_{j=1}^m \left(\int_{\mathbb{R}^{n_j}} f_j \right)^{c_j} \tag{6.8}$$

holds true for all $f_j \in L_1^+(\mathbb{R}^{n_j})$. Then, we have the following theorem.

Theorem 2.6.1 *The constants E and F can be computed using centered Gaussian functions. That is,*

$$E = \inf\left\{ \frac{K(g_1,\dots,g_m)}{\prod_{j=1}^m \left(\int_{\mathbb{R}^{n_j}} g_j\right)^{c_j}} \mid g_j \ \textit{ is a centered Gaussian }, \ j=1,\dots,m\right\} \tag{6.9}$$

and

$$F = \sup\left\{ \frac{I(g_1,\dots,g_m)}{\prod_{j=1}^m \left(\int_{\mathbb{R}^{n_j}} g_j\right)^{c_j}} \mid g_j \ \textit{ is a centered Gaussian }, \ j=1,\dots,m\right\}. \tag{6.10}$$

Moreover, if D is the largest real number for which

$$\det\left(\sum_{j=1}^m c_j B_j^* A_j B_j\right) \geq D \cdot \prod_{j=1}^m (\det A_j)^{c_j}, \tag{6.11}$$

for all $A_j \in S^+(\mathbb{R}^{n_j})$, we have

$$E = \sqrt{D} \quad \textit{and} \quad F = \frac{1}{\sqrt{D}}. \tag{6.12}$$

The proof follows the steps of the one-dimensional case. We consider the constants

$$E_g = \inf\left\{ \frac{K(g_1,\dots,g_m)}{\prod_{j=1}^m \left(\int_{\mathbb{R}^{n_j}} g_j\right)^{c_j}} \mid g_j \ \text{ is a centered Gaussian }, \ j=1,\dots,m\right\} \tag{6.13}$$

and

$$F_g = \sup\left\{ \frac{I(g_1,\dots,g_m)}{\prod_{j=1}^m \left(\int_{\mathbb{R}^{n_j}} g_j\right)^{c_j}} \mid g_j \ \text{ is a centered Gaussian }, \ j=1,\dots,m\right\}. \tag{6.14}$$

What we want to prove is that

$$E = E_g = \sqrt{D} \quad \text{and} \quad F = F_g = \frac{1}{\sqrt{D}}. \tag{6.15}$$

Lemma 2.6.1 *In the notation of Theorem 2.6.1, we have $F_g = 1/\sqrt{D}$.*

Proof: Let $g_j = G_{A_j}$, $j=1,\dots,m$, where $A_j \in S^+(\mathbb{R}^{n_j})$. Then,

$$\begin{aligned}
\int_{\mathbb{R}^n} \prod_{j=1}^m g_j^{c_j}(B_j x)dx &= \int_{\mathbb{R}^n} \exp\left(-\sum_{j=1}^m c_j \langle A_j B_j x, B_j x\rangle\right) dx \\
&= \int_{\mathbb{R}^n} \exp\left(-\langle\big(\sum_{j=1}^m c_j B_j^* A_j B_j\big)(x), x\rangle\right) dx \\
&= \frac{\pi^{n/2}}{\sqrt{\det\left(\sum_{j=1}^m c_j B_j^* A_j B_j\right)}}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\prod_{j=1}^{m}\left(\int_{\mathbb{R}} G_{A_j}\right)^{c_j} &= \prod_{j=1}^{m}\left(\int_{\mathbb{R}} \exp(-\langle A_j x, x\rangle)\right)^{c_j} \\
&= \prod_{j=1}^{m}\left(\frac{\pi^{n_j/2}}{\sqrt{\det A_j}}\right)^{c_j} \\
&= \frac{\pi^{n/2}}{\sqrt{\prod_{j=1}^{m}(\det A_j)^{c_j}}}
\end{aligned}$$

since $c_1 n_1 + \cdots + c_m n_m = n$. It follows that

$$\begin{aligned}
\frac{1}{F_g^2} &= \inf\left\{\left(\frac{\prod_{j=1}^{m}\left(\int_{\mathbb{R}^{n_j}} g_j\right)^{c_j}}{\int_{\mathbb{R}^n}\prod_{j=1}^{m} g_j^{c_j}(B_j x)dx}\right)^2 \mid g_j \text{ is a centered Gaussian },\ j=1,\dots,m\right\} \\
&= \inf\left\{\frac{\det\left(\sum_{j=1}^{m} c_j B_j^* A_j B_j\right)}{\prod_{j=1}^{m}(\det A_j)^{c_j}} \mid A_j \in S^+(\mathbb{R}^{n_j})\right\} \\
&= D. \qquad \square
\end{aligned}$$

Lemma 2.6.2 *We have $E_g \cdot F_g = 1$ and $E_g = 0$ if and only if $F_g = +\infty$. Therefore, $F_g = \sqrt{D}$.*

Proof: Let $A_j \in S^+(\mathbb{R}^{n_j})$ and consider the quadratic form Q defined by

$$Q(y) = \Big\langle \sum_{j=1}^{m} c_j B_j^* A_j B_j y, y\Big\rangle. \tag{6.16}$$

We consider the function

$$R(x) = \inf\left\{\sum_{j=1}^{m} c_j\langle A_j^{-1}x_j, x_j\rangle \mid x_j \in \mathbb{R}^{n_j} \text{ and } x = \sum_{j=1}^{m} c_j B_j^* x_j\right\}. \tag{6.17}$$

We will show that

$$R(x) = Q^*(x) = \sup\{\langle x, y\rangle^2 \mid Q(y) \le 1\}. \tag{6.18}$$

Note that if $x = \sum_{j=1}^{m} c_j B_j^* x_j$ where $x_j \in \mathbb{R}^{n_j}$, then

$$\langle x, y\rangle^2 = \Big\langle \sum_{j=1}^{m} c_j B_j^* x_j, y\Big\rangle^2 = \left(\sum_{j=1}^{m}\langle\sqrt{c_j}A_j^{-1/2}x_j, \sqrt{c_j}A_j^{1/2}B_j y\rangle\right)^2,$$

and hence, Cauchy-Schwarz inequality shows that

$$\langle x, y\rangle^2 \le \left(\sum_{j=1}^{m}|\sqrt{c_j}A_j^{-1/2}x_j|^2\right)\cdot\left(\sum_{j=1}^{m}|\sqrt{c_j}A_j^{1/2}B_j y|^2\right)$$

$$
\begin{aligned}
&= \left(\sum_{j=1}^{m} c_j \langle x_j, A_j^{-1} x_j \rangle\right) \cdot \left(\langle \sum_{j=1}^{m} c_j B_j^* A_j B_j y, y\rangle\right) \\
&\leq R(x)Q(y).
\end{aligned}
$$

On the other hand, we have equality in the previous computation if we choose

$$
y = \left(\sum_{j=1}^{m} c_j B_j^* A_j B_j\right)^{-1} (x)
$$

and $x_j = A_j B_j y$, $j = 1, \ldots, m$. It follows that $R = Q^*$.

Direct computation shows that if $A_j \in S^+(\mathbb{R}^{n_j})$, then

$$
\frac{I(G_{A_1}, \ldots, G_{A_m})}{\prod_{j=1}^{m} \left(\int G_{A_j}\right)^{c_j}} = \left(\frac{\prod_{j=1}^{m} (\det A_j)^{c_j}}{\det Q}\right)^{1/2} \tag{6.19}
$$

and

$$
\frac{K(G_{A_1^{-1}}, \ldots, G_{A_m^{-1}})}{\prod_{j=1}^{m} \left(\int G_{A_j^{-1}}\right)^{c_j}} = \left(\frac{\prod_{j=1}^{m} (\det A_j)^{-c_j}}{\det R}\right)^{1/2}. \tag{6.20}
$$

Since $R = Q^*$, we have $(\det Q) \cdot (\det R) = 1$. Therefore,

$$
\frac{I(G_{A_1}, \ldots, G_{A_m})}{\prod_{j=1}^{m} \left(\int G_{A_j}\right)^{c_j}} \cdot \frac{K(G_{A_1^{-1}}, \ldots, G_{A_m^{-1}})}{\prod_{j=1}^{m} \left(\int G_{A_j^{-1}}\right)^{c_j}} = 1. \tag{6.21}
$$

By the definition of E_g and F_g we get the conclusion of the lemma. □

The main step for the proof of Theorem 2.6.1 is the next proposition. We may clearly assume that $D > 0$.

Proposition 2.6.1 *Assume that $h_j, f_j \in L_1^+(\mathbb{R}^{n_j})$, $j = 1, \ldots m$ satisfy*

$$
\int_{\mathbb{R}^{n_j}} h_j = \int_{\mathbb{R}^{n_j}} f_j = 1.
$$

Then,

$$
K(h_1, \ldots, h_m) \geq D \cdot I(f_1, \ldots, f_m).
$$

Proof: We denote by $\mathcal{C}_L(\mathbb{R}^k)$ the class of functions $f \in L_1^+(\mathbb{R}^k)$ which are restrictions of some positive Lipschitz function $F : \mathbb{R}^k \to \mathbb{R}$ onto some open Euclidean ball. It suffices to prove the proposition for functions $f_j, h_j \in \mathcal{C}_L(\mathbb{R}^{n_j})$.

We write $\Omega_{f_j}, \Omega_{h_j}$ for the balls $\{f_j > 0\}, \{h_j > 0\}$. We will use the Brenier map: there exists a function $T_j : \Omega_{f_j} \to \Omega_{h_j}$ such that

$$
|J(T_j)(x)| \cdot h_j(T_j x) = f_j(x) \tag{6.22}
$$

for all $x \in \Omega_{f_j}$. Moreover T_j is the gradient of a convex function and a regularity result of Caffarelli shows that T_j is continuously differentiable. It follows that $J(T_j)(x)$ is a symmetric positive definite matrix for every $x \in \Omega_{f_i}$.

We write $S = \bigcap_{j=1}^m B_j^{-1}(\Omega_{f_j}) \subseteq \mathbb{R}^n$ and define a map $\Theta : S \to \mathbb{R}^n$ by

$$\Theta(y) = \sum_{j=1}^{m} c_j B_j^*(T_j(B_j(y))). \tag{6.23}$$

Direct computation shows that

$$J(\Theta)(y) = \sum_{j=1}^{m} c_j B_j^*\big(J(T_j)(B_j y)\big)B_j. \tag{6.24}$$

Note that

$$|J(\Theta)(y)| \geq D \cdot \prod_{j=1}^{m} |J(T_j)(B_j y)|^{c_j} > 0, \tag{6.25}$$

and hence, $J(\Theta)(y)$ is symmetric positive definite. It follows that Θ is injective. Using (6.22) and (6.25) we write

$$\begin{aligned}
K(h_1,\ldots,h_m) &= \int_{\mathbb{R}^n} \sup\left\{ \prod_{j=1}^m h_j^{c_j}(B_j^* x_j) \mid x = \sum_{j=1}^m c_j B_j^* x_j \right\} dx \\
&\geq \int_{\Theta(S)} \sup\left\{ \prod_{j=1}^m h_j^{c_j}(B_j^* x_j) \mid x = \sum_{j=1}^m c_j B_j^* x_j \right\} dx \\
&= \int_S |J(\Theta)(y)| \cdot \sup\left\{ \prod_{j=1}^m h_j^{c_j}(B_j^* y_j) \mid \Theta(y) = \sum_{j=1}^m c_j B_j^* y_j \right\} dy \\
&\geq \int_S |J(\Theta)(y)| \cdot \prod_{j=1}^m h_j^{c_j}(T_j(B_j y)) dy \\
&\geq D \cdot \int_S \prod_{j=1}^m |J(T_j)(B_j y)|^{c_j} \prod_{j=1}^m h_j^{c_j}(T_j(B_j y)) dy \\
&= D \cdot \int_S \prod_{j=1}^m f_j^{c_j}(B_j y) dy \\
&= D \cdot I(f_1,\ldots,f_m). \qquad \square
\end{aligned}$$

Proof of Theorem 2.6.1: Combining Lemma 2.6.1, Lemma 2.6.2 and Proposition 2.6.1, we have

$$\frac{1}{\sqrt{D}} = F_g \quad \leq \quad \sup\left\{ I(f_1,\ldots,f_m) \mid \int_{\mathbb{R}} f_j = 1 \right\}$$

$$\begin{aligned}
&\leq \frac{1}{D} \cdot \inf \left\{ K(h_1, \ldots, h_m) \mid \int_{\mathbb{R}} h_j = 1 \right\} \\
&\leq E_g \cdot \frac{1}{D} = \frac{1}{\sqrt{D}}.
\end{aligned}$$

We must have equality everywhere, and this proves the theorem. □