Inequalities in convex geometry

Spring semester, 2017/2018

Piotr Nayar, nayar@mimuw.edu.pl Institute of Mathematics, University of Warsaw

1. Brunn-Minkowski inequalities

1.1. Classical BM inequality. To avoid problems with measurability we assume that K, Lare compact sets. Then their Minkowski sum

$$K + L = \{a + b : a \in K, b \in L\}$$

is also compact. The Brunn-Minkowski inequality gives a lower bound on the volume of K + L, namely for **non-empty** compact sets K, L in \mathbb{R}^n we have

(1)
$$|K+L|^{1/n} \ge |K|^{1/n} + |L|^{1/n}.$$

Here $|\cdot|$ stand for the *n*-dimensional Lebesgue measure, which will be sometimes denoted by vol_n , to emphasise the dependence of n. This inequality is equivalent to its multiplicative form, which seems to be weaker,

(2)
$$|\lambda K + (1-\lambda)L| \ge |K|^{\lambda} |L|^{1-\lambda}, \qquad \lambda \in [0,1]$$

By considering λK instead of K and $(1 - \lambda)L$ instead of L we see that (1) is equivalent to

(3)
$$|\lambda K + (1-\lambda)L|^{1/n} \ge \lambda |K|^{1/n} + (1-\lambda)|L|^{1/n}.$$

Clearly (3) implies (2) by applying AM-GM inequality $\lambda a + (1 - \lambda)b \ge a^{\lambda}b^{1-\lambda}$ to the right hand side of (3). There is an even weaker formulation of BM inequality, namely

(4)
$$|K| = |L| = 1 \implies |\lambda K + (1 - \lambda)L| \ge 1, \quad \lambda \in [0, 1].$$

Clearly (2) implies (4). To see that (4) implies (3) we apply (4) with $\tilde{K} = K/|K|^{1/n}$, $\tilde{L} = L/|L|^{1/n}$ and

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

Thus, (1), (2), (4) and (3) are all equivalent.

Another equivalent way to state BM inequality is to say that for any two sets K, L if we take two balls B_K, B_L such that $|K| = |B_K|$ and $|L| = |B_L|$ then

$$|K+L| \ge |B_K+B_L|.$$

Indeed, if $B_K = r_K B_2^n$ and $B_L = r_L B_2^n$ then

$$|B_K + B_L| = (r_K + r_L)^n |B_2^n| = \left(\left(\frac{|K|}{B_2^n}\right)^{1/n} + \left(\frac{|L|}{B_2^n}\right)^{1/n} \right)^n |B_2^n| = (|K|^{1/n} + |L|^{1/n})^n.$$

1.2. **Proof of BM via elementary sets.** Let us first show that in order to prove (1) for measurable set A, B such that the sum is measurable, it suffices to consider only compact sets. Indeed, by the regularity of Lebesgue measure we can approximate them form below by compact sets $A_{\varepsilon}, B_{\varepsilon}$ and write

(5)
$$|A+B|^{1/n} \ge |A_{\varepsilon}+B_{\varepsilon}|^{1/n} \ge |A_{\varepsilon}|^{1/n} + |B_{\varepsilon}|^{1/n} \xrightarrow[\varepsilon \to 0^+]{} |A|^{1/n} + |B|^{1/n}$$

Assume that we could prove the inequality for open sets. Then compact sets A, B it suffices to take open ε enlargements $A_{\varepsilon}, B_{\varepsilon}$ and write

$$|A + B + 2\varepsilon B_2^n|^{1/n} = |A_{\varepsilon} + B_{\varepsilon}|^{1/n} \ge |A_{\varepsilon}|^{1/n} + |B_{\varepsilon}|^{1/n} \ge |A|^{1/n} + |B|^{1/n}.$$

We conclude by taking $\varepsilon \to 0^+$ and using the continuity of the measure. Now, every open set can be approximated from below by a finite union of boxes of the form

$$v + [-a_1, a_1] \times \ldots \times [-a_n, a_n], \qquad v \in \mathbb{R}^n, \ a_1, \ldots a_n > 0$$

If we could prove our inequality for such sets an argument identical to (5) would finish the proof.

So, it suffices to consider only finite unions of boxes. If both A and B are just two boxes (note that by translation invariance of Lebesgue measure we can assume that these boxes are centred at 0)

$$A = [-a_1, a_1] \times \ldots \times [-a_n, a_n], \qquad B = [-b_1, b_1] \times \ldots \times [-b_n, b_n]$$

then we are to verify

$$\sqrt[n]{(a_1+b_1)\dots(a_n+b_n)} \ge \sqrt[n]{a_1\dots a_n} + \sqrt[n]{b_1\dots b_n}.$$

Note that by AM-GM we get

$$\sqrt[n]{\frac{a_1}{a_1+b_1}\cdot\ldots\cdot\frac{a_n}{a_n+b_n}} \le \frac{1}{n}\left(\frac{a_1}{a_1+b_1}+\ldots+\frac{a_n}{a_n+b_n}\right)$$

and

$$\sqrt[n]{\frac{b_1}{a_1+b_1}\cdot\ldots\cdot\frac{b_n}{a_n+b_n}} \leq \frac{1}{n}\left(\frac{b_1}{a_1+b_1}+\ldots+\frac{b_n}{a_n+b_n}\right).$$

Adding these two proves our inequality.

Now we use induction on the number m of boxes used in the union of A and B. First, we find a hyperplane of the form $x_i = s$ for some i = 1, ..., n and $s \in \mathbb{R}$ such that the sets $A^+ = A \cap \{x_i \ge s\}$ and $A^- = A \cap \{x_i \le s\}$ both consist of a smaller number of boxes that the original set A (it is easy to see that one can always find a cut separating at least two boxes). Now, find a number $t \in \mathbb{R}$ such that

$$\frac{|A_+|}{|A|} = \frac{|B_+|}{|B|} =: \lambda,$$

where $B_+ = B \cap \{x_i \ge t\}$ and $B_- = B \cap \{x_i \le t\}$. The sets $A_+ \cup B_+$ and $A_- \cup B_-$ are disjoint and both consist of a number of boxes smaller than m. By induction hypothesis

$$|A + B| \ge |A_{+} + B_{+}| + |A_{-} + B_{-}| \ge (|A_{+}|^{1/n} + |B_{+}|^{1/n})^{n} + (|A_{-}|^{1/n} + |B_{-}|^{1/n})^{n} \ge \lambda (|A|^{1/n} + |B|^{1/n})^{n} + (1 - \lambda)(|A|^{1/n} + |B|^{1/n})^{n} = (|A|^{1/n} + |B|^{1/n})^{n}.$$

1.3. Brunn's principle. We shall prove the following theorem.

Theorem 1. Suppose K is a convex body in \mathbb{R}^n and let $u \in S^{n-1}$. Then the function

$$t \mapsto \operatorname{vol}_{n-1}(K \cap (u^{\perp} + tu))^{1/n-1}$$

is concave on its support.

Proof. We can assume that $u = e_1$. Let $K_t = K \cap (u^{\perp} + tu) = K \cap \{x_1 = t\}$ and consider these as sets in \mathbb{R}^{n-1} . We claim that $\lambda K_t + (1-\lambda)K_s \subseteq K_{\lambda t+(1-\lambda)s}$. Indeed, suppose $a \in K_t$ and $b \in K_s$. Then by convexity of K we have $\lambda(t, a) + (1-\lambda)(s, b) = (\lambda t + (1-\lambda)s, \lambda a + (1-\lambda)b) \in$ K and thus $\lambda a + (1-\lambda)b \in K_{\lambda t+(1-\lambda)s}$. Suppose K_s, K_t are non-empty (i.e. we are on the support of our map). By Brunn-Minkowski we get

$$|K_{\lambda t+(1-\lambda)s}|^{\frac{1}{n-1}} \ge |\lambda K_t + (1-\lambda)K_s|^{\frac{1}{n-1}} \ge \lambda |K_t|^{\frac{1}{n-1}} + (1-\lambda)|K_s|^{\frac{1}{n-1}},$$

which proves the desired concavity.

1.4. Isoperimetric inequality. For a compact sets K in \mathbb{R}^n we define $K_t = K + tB_2^n$.

Theorem 2. Let K be a compact set in \mathbb{R}^n and let B be a ball such that |K| = |B|. Then

(a)
$$|K_t| \ge |B_t| = \left(\left(\frac{|K|}{|B_2^n|} \right)^{1/n} + t \right)^n |B_2^n|$$

(b) $|\partial K| \ge |\partial B| = n|K|^{\frac{n-1}{n}} |B_2^n|^{\frac{1}{n}}.$

Proof. Suppose $B = rB_2^n$. By the BM-inequality we have

$$|K_t| = |K + tB_2^n| \ge \left(|K|^{\frac{1}{n}} + t|B_2^n|^{\frac{1}{n}}\right)^n = \left(|B|^{\frac{1}{n}} + t|B_2^n|^{\frac{1}{n}}\right)^n$$
$$= (r+t)^n |B_2^n| = |(r+t)B_2^n| = |B + tB_2^n| = |B_t|.$$

To prove the second part we recall that

$$|\partial K| = \liminf_{\varepsilon \to 0^+} \frac{|K + \varepsilon B_2^n| - |K|}{\varepsilon}$$

Thus from point (a) we get

$$|\partial K| = \liminf_{\varepsilon \to 0^+} \frac{|K_{\varepsilon}| - |K|}{\varepsilon} \ge \liminf_{\varepsilon \to 0^+} \frac{|B_{\varepsilon}| - |B|}{\varepsilon} = |\partial B|.$$

1.5. Steiner symmetrization. In this section we will usually assume that the sets K, L are convex. For a measurable set A and a unite vector u in \mathbb{R}^n we define Steiner symmetral by

$$S_u A = \left\{ (x, tu) : x \in \operatorname{Proj} u^{\perp}(A), |t| \leq \frac{1}{2} \operatorname{vol}_1(A \cap (x + u\mathbb{R})) \right\}.$$

In other words, for every line l perpendicular to u^{\perp} we replace the intersection of A with this line with an interval symmetric with respect to u^{\perp} , of the same 1-dimensional measure as $l \cap A$.

We shall need several useful properties of Steiner symmetrization. In the below theorem the volume of the boundary of K is defined via

$$|\partial K| = \liminf_{\varepsilon \to 0^+} \frac{|K + \varepsilon B_2^n| - |K|}{\varepsilon}.$$

Proposition 3. Let K, L be convex compact sets in \mathbb{R}^n . Then

- (i) $S_u(\lambda K) = \lambda S_u K$, (ii) $K \subseteq L \implies S_u K \subseteq S_u L$, (iii) $S_u K$ is convex and compact, (iv) $|S_u K| = |K|$, (vi) $S_u K + S_u L \subseteq S_u(K + L)$,
- (vii) $|\partial S_u K| \leq |\partial K|$.

Proof. Points (i), (ii) are very easy. Point (iv) follows from Fubini (Cavalieri's principle).

(iii) Let $x_1, x_2 \in S_u K$ and let $l_{x_i} = \operatorname{Proj}_{u^{\perp}} x_i + u\mathbb{R}$, i = 1, 2. The convex hull of line segments $l_{x_i} \cap S_u K$ is a two-dimensional trapezoid T. Since $x, y \in T$ and T is convex, to prove that $\lambda x + (1 - \lambda)y \in S_u K$, it suffices to show that $T \subseteq S_u K$. This follows from the fact that T is the Steiner symmetral of the trapezoid \tilde{T} defined as the convex hull of $l_{x_i} \cap K$. Since $\tilde{T} \subseteq K$ we get by (ii) that $T = S_u \tilde{T} \subseteq S_u K$. Compactness of $s_u K$ follows from its convexity and from the fact that for any x the line l_x is a closed interval (we shall skip the details). The boundedness of $S_u K$ follows from the boundedness of K since by (ii) if $K \subseteq B$, where B is a centred Euclidean ball, the $S_u K \subseteq S_u B = S_u B$.

(vi) Let $x \in S_u K$ and $y \in S_u L$. We shall prove that $x + y \in S_u(K + L)$. We have x = (x', tu), y = (y', su), where $x', y' \in u^{\perp}$ and $|t| \leq \frac{1}{2}|K \cap l_x|, |s| \leq \frac{1}{2}|K \cap l_y|$. We have x + y = (x' + y', u(s + t)). Therefore, it suffices to show that $|s + t| \leq \frac{1}{2}|(K + L) \cap l_{x+y}|$. In fact it is enough to show that $|(K + L) \cap l_{x+y}| \geq |K \cap l_x| + |L \cap l_x|$. This is true since

$$K \cap (x + \mathbb{R}u) + L \cap (y + \mathbb{R}u) \subseteq (K + L) \cap (x + y + \mathbb{R}u)$$

and the left hand side is an interval of length $|K \cap l_x| + |L \cap l_x|$.

(vii) We have (using (i), (iv) and (vi)) that

$$\begin{aligned} |\partial S_u K| &= \liminf_{\varepsilon \to 0^+} \frac{|S_u K + \varepsilon B_2^n| - |S_u K|}{\varepsilon} = \liminf_{\varepsilon \to 0^+} \frac{|S_u K + \varepsilon S_u B_2^n| - |K|}{\varepsilon} \\ &\leq \liminf_{\varepsilon \to 0^+} \frac{|S_u (K + \varepsilon B_2^n)| - |K|}{\varepsilon} = \liminf_{\varepsilon \to 0^+} \frac{|K + \varepsilon B_2^n| - |K|}{\varepsilon} = |\partial K|. \end{aligned}$$

The Hausdorff distance between convex bodies is defined by

$$d_H(K,L) = \inf\{\delta > 0: K \subseteq L + \delta B_2^n, L \subseteq K + \delta B_2^n\}.$$

It is not hard to see that

$$d_H(K, L) = \sup_{u \in S^{n-1}} |h_K(u) - h_L(u)|,$$

where $h_K(u) = \sup\{\langle x, u \rangle : x \in K\}$ is the so-called support function. Indeed, this follows from the fact that $h_{K+L} = h_K + h_L$ and from the facts that $K \subseteq L$ is equivalent to $h_K \leq h_L$. We will also need the following fact.

Proposition 4. For convex sets in \mathbb{R}^n with non-empty interior the following holds true.

- (i) If $K_i \to K$ then $|K_i| \to |K|$.
- (ii) If $K_j \to K$ and $L_j \to L$ then $K_j + L_j \to K + L$.
- (iii) If $d_H(K, B) \leq \varepsilon$ for certain ball B then $d_H(S_uK, B) \leq \varepsilon$.

Proof. (i) If K_j and K are ε -close, that is for any $u \in S^{n-1}$ we have $|h_{K_j}(u) - h_K(u)| \leq \varepsilon$ then $|h_{K_j}(u) - h_K(u)| \leq C \varepsilon h_K(u)$, where $C = 1/\inf_u h_K(u)$. Thus $(1 - C\varepsilon)K \subseteq K_j \subseteq (1 + C\varepsilon)K$ and the convergence of volumes follows in the limit $\varepsilon \to 0^+$.

(ii) This is clear due to the relation $h_{K_j+L_j} = h_{K_j} + h_{L_j}$ and the fact that $K_j \to K$ if and only if $h_{K_j} \to h_K$ uniformly.

(iii) If B has radius r then $d_H(K, B) \leq \varepsilon$ is equivalent to $|h_K - r| \leq \varepsilon$ which is $r - \varepsilon \leq h_K \leq r + \varepsilon$, that is $(r - \varepsilon)B_2^n \subseteq K \subseteq (r + \varepsilon)B_2^n$. The latter is preserved under $K \to S_u K$. \Box

Theorem 5 (Blaschke selection principle). Any sequence of convex bodies $(K_j)_{j=1}^{\infty}$ in \mathbb{R}^n of which all elements are contained in some fixed ball, has a convergent subsequence.

Proof. Step 1. We construct an array of bodies $(K_{i,j})_{i,j=1}^{\infty}$, all of which belong to our original sequence, such that $(K_{i,j})_{j=1}^{\infty}$ is a subsequence of $(K_{i-1,j})_{j=1}^{\infty}$ and $d_H(K_{i,j_1}, K_{i,j_2}) \leq 2^{-i}$ for any j_1, j_2 . To do this at each step we coved the big ball (which contains our sequence) by a finite number of balls of diameter 2^{-i} . Let us call these balls B_1, \ldots, B_{N_i} . There is an infinite subsequence such that either all elements intersect B_1 or all elements do not intersect B_1 . By passing to a further subsequence $N_i - 1$ times we get a subsequence such that for any ball B_j either all elements intersect this ball, or neither of them intersect it. Suppose there is a point x in K_{i,j_1} such that $d(K_{i,j_2}, x) > 2^{-i}$. Then the ball B_j covering x do not intersect K_{i,j_1} , which is a contradiction with our construction. Thus, $d_H(K_{i,j_1}, K_{i,j_2}) \leq 2^{-i}$.

Step 2. Take the diagonal $(K_{j,j})_{j=1}^{\infty}$ to get $d(K_{j,j}, K_{i,i}) \leq 2^{-\min(i,j)}$. Assume we have $K_{j,j} = K_{n_j}$ in the numbering of the original sequence K_1, K_2, \ldots . We claim that $K_{n_j} + \frac{1}{2^{j-1}}B_2^n$ is monotone decreasing. Indeed, $K_{n_{j+1}} \subseteq K_{n_j} + \frac{1}{2^j}B_2^n$ and thus

$$K_{n_{j+1}} + \frac{1}{2^j} B_2^n \subseteq K_{n_j} + \frac{1}{2^j} B_2^n + \frac{1}{2^j} B_2^n = K_{n_j} + \frac{1}{2^{j-1}} B_2^n.$$

Step 3. We take $K = \bigcap_j (K_{n_j} + \frac{1}{2^{j-1}}B_2^n)$. Note that K is clearly convex. We claim K is the limit of (K_{n_j}) . For j large enough so that $2^{-(j-1)} < \varepsilon$ we have $K \subseteq K_{n_j} + \varepsilon B_2^n$. Take $G = \operatorname{int}(K + \varepsilon B_2^n)$. We have that $(K_{n_j} + \frac{1}{2^{j-1}}B_2^n) \setminus G$ are compact and

$$\bigcap_{j=1}^{\infty} (K_{n_j} + \frac{1}{2^{j-1}} B_2^n) \setminus G = K \cap (\mathbb{R}^n \setminus G) = \emptyset.$$

Since the intersection of decreasing family of non-empty compact sets is non-empty, from some point onwards we must have

$$K_{n_j} \subseteq K_{n_j} + \frac{1}{2^{j-1}} B_2^n \subseteq G \subseteq K + \varepsilon B_2^n.$$

Since $K \subseteq K_{n_j} + \varepsilon B_2^n$ and $K_{n_j} \subseteq K + \varepsilon B_2^n$ we get $d_H(K, K_{n_j}) \leq \varepsilon$ for large j.

Proposition 6. For any K_1, \ldots, K_m there is a sequence of vectors u_1, u_2, \ldots such that

$$S_{u_k} \dots S_{u_2} S_{u_1} K_i \to_{k \to \infty} \left(\frac{|K|}{|B_2^n|} \right)^{1/n} B_2^n, \qquad i = 1, \dots, m$$

Proof. It is enough to prove the claim only for one convex body, since due to (iii) once some of the bodies are already close to a ball, they will stay close after applying arbitrary symmetrizations. We then apply standard diagonal argument.

We consider the class of bodies \mathcal{K} obtained by successive symmetrizations of K. If $K \subseteq rB_2^n$ for some r, then all the members of \mathcal{K} are contained in rB_2^n . Let

$$R_0 = \inf\{\operatorname{circumradius}(\mathbf{K}) : \mathbf{K} \in \mathcal{K}\}.$$

Take a sequence in \mathcal{K} with circumradi converging to R_0 and by Blaschke selection principle pass to a subsequence such that $K_j \to L$ for some convex body L. It is easy to see (by the definition of Hausdorff distance) that R_0 is the circumradius of L. We claim that $L = R_0 B_2^n$. Suppose ∂L misses some cup

$$C = \partial(R_0 B_2^n) \cap \{ \langle x, y \rangle \ge R_0 - \varepsilon \}, \qquad |u| = 1.$$

It is not hard to see that there is a sequence of hyperplanes H_1, \ldots, H_k such that

$$\partial(R_0 B_2^n) = \bigcup_{i=1}^k S_{H_i}(C).$$

Indeed if we want to cover some point $x \in \partial(R_0B_2^n)$ by a mirror image of C, it suffices to use hyperplane H_x perpendicular to $x - x_0$, where x_0 in the center of C. We then choose a finite subcovering by compactness. If L misses the cap C then $S_{H_x}(L)$ misses both C and $S_{H_x}(C)$. Thus $L_0 = S(L)$ where $S = S_{H_k} \circ \ldots \circ S_{H_1}$ misses all the $\partial(R_0B_2^n)$. Since L_0 is compact we have $L_0 \subseteq (R_0 - \varepsilon)B_2^n$ for some $\varepsilon > 0$. Suppose $B_2^n \subseteq t_0L_0$ and define

$$\tilde{\varepsilon} = \left(\frac{R_0 - \varepsilon/2}{R_0 - \varepsilon} - 1\right) \frac{1}{t_0}.$$

Suppose $L_1 \in \mathcal{K}$ is $\tilde{\varepsilon}$ -close to L_0 . Then

$$L_1 \subseteq L_0 + \tilde{\varepsilon}B_2^n \subseteq L_0 + \tilde{\varepsilon}t_0L_0 = (1 + \tilde{\varepsilon}t_0)L_0 \subseteq (1 + \tilde{\varepsilon}t_0)(R_0 - \varepsilon)B_2^n = (R_0 - \varepsilon/2)B_2^n$$

This contradicts the definition of R_0 .

We are ready to give a proof of the BM inequality for convex sets.

Proof of the BM inequality. Take u_1, u_2, \ldots a sequence given by Proposition 4(iv) for $K_1 = K$ and $K_2 = L$. From Proposition 3(vi) and (iv) we get that

$$|S_{u_k} \dots S_{u_2} S_{u_1} K + S_{u_k} \dots S_{u_2} S_{u_1} L| \le |K + L|,$$

whereas $|S_{u_k} \dots S_{u_2} S_{u_1} K| = |K|$ and $|S_{u_k} \dots S_{u_2} S_{u_1} L| = |L|$. From Proposition 4(i) and (ii) we infer that

$$\left|S_{u_k}\dots S_{u_2}S_{u_1}K + S_{u_k}\dots S_{u_2}S_{u_1}L\right| \to \left|\left(\frac{|K|}{|B_2^n|}\right)^{1/n}B_2^n + \left(\frac{|L|}{|B_2^n|}\right)^{1/n}B_2^n\right| = (|K|^{1/n} + |L|^{1/n})^n.$$

Thus,

$$(|K|^{1/n} + |L|^{1/n})^n \le |K + L|.$$

1.6. Applications of Steiner symmetrization and BM inequality.

Urysohn's inequality. Let K be a convex body in \mathbb{R}^n . We define the support function $h_K: S^{n-1} \to \mathbb{R}$ via

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

The mean width is defined via the formula

$$\omega(K) = \int_{S^{n-1}} h_K(u) \mathrm{d}\sigma(u),$$

where σ in the uniform probability measure on S^{n-1} .

Theorem 7. For any convex body in \mathbb{R}^n we have

$$\omega(K) \ge \left(\frac{|K|}{|B_2^n|}\right)^{1/n}.$$

Clearly in the above inequality we have equality for centered Euclidean balls. Since Steiner symmetrization does not change volume due to Proposition 6, it suffices to prove the following lemma.

Lemma 1. For any convex body K in \mathbb{R}^n and any $\theta \in S^{n-1}$ we have $\omega(S_{\theta}(K)) \leq \omega(K)$. *Proof.* Without loss of generality we assume $\theta = e_n$. Then

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

This is due to the fact that Steiner symmetrization of an interval I on the real line is equal to $\frac{1}{2}(I-I)$. For $u = (u_1, \ldots, u_n) \in S^{n-1}$ we take $u' = (u_1, \ldots, u_{n-1}, -u_n)$. then

$$h_{S_{\theta}(K)}(u) = \max\left\{\left\langle \left(x, \frac{t_1 - t_2}{2}\right), u\right\rangle : (x, t_1), (x, t_2) \in K\right\} \\ \leq \frac{1}{2} \max\left\{\left\langle (x, t_1), u\right\rangle : (x, t_1) \in K\right\} + \frac{1}{2} \max\left\{\left\langle (x, t_2), u'\right\rangle : (x, t_2) \in K\right\} \\ = \frac{1}{2} h_K(u) + \frac{1}{2} h_K(u').$$

Thus,

$$\omega(S_{\theta}(K)) = \int_{S^{n-1}} h_{S_{\theta}(K)}(u) \mathrm{d}\sigma(u) \leq \frac{1}{2} \int_{S^{n-1}} h_K(u) \mathrm{d}\sigma(u) + \frac{1}{2} \int_{S^{n-1}} h_K(u') \mathrm{d}\sigma(u)$$
$$= \int_{S^{n-1}} h_K(u) \mathrm{d}\sigma(u) = \omega(K).$$

Blashke-Santalo inequality. For a compact set K in \mathbb{R}^n we define

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

We prove the following theorem.

Theorem 8. If K is a centrally symmetric convex body in \mathbb{R}^n then

$$|K| \cdot |K^{\circ}| \le |B_2^n|^2.$$

Again it suffices to prove monotonicity under Steiner symmetrization.

Lemma 2. For any centrally symmetric convex body K and any $\theta \in S^{n-1}$ we have $|K^{\circ}| \leq |S_{\theta}(K)^{\circ}|$.

Indeed if we denote $v(K) = |K| \cdot |K^{\circ}|$ then $v(K) \leq v(S_uK)$ and due to Proposition 6 we can find a sequence of directions u_1, \ldots, u_n such that $(S_{u_n} \circ \ldots \circ S_{u_1}(K))$ converges to a certain ball B and thus

$$v(K) \le v(S_{u_n} \circ \ldots \circ S_{u_1}(K)) \to v(B) = v(B_2^n)$$

Note that we have used the fact that the convergence $K_n \to B$ implies $K_n^{\circ} \to B^{\circ}$ which can we easily seen by dualizing the inclusion $(r - \varepsilon)B_2^n \subseteq K_n \subseteq (r - \varepsilon)B_2^n$ (see the proof

of Proposition 4(iii)). Note also that the last equality is a consequence of the fact that v is linear invariant, which follows from the fact that for any invertible linear map T we have $T(A)^{\circ} = ((T^*)^{-1})(A^{\circ})$ (exercise) and thus

$$v(TK) = |TK| \cdot |(TK)^{\circ}| = |\det(T)||\det(T^{*})^{-1}| \cdot |K| \cdot |K^{\circ}| = |K| \cdot |K^{\circ}| = v(K).$$

Proof of Lemma 2. Without loss of generality we can assume that $\theta = e_n$. Again we have

$$S_{\theta}(K) = \left\{ \left(x, \frac{s-t}{2}\right) : (x,s), (x,t) \in K \right\}.$$

We get

$$(S_{\theta}K)^{\circ} = \{(y,r): \langle x,y \rangle + \frac{1}{2}r(s-t) \le 1, (x,s), (x,t) \in K\}.$$

Define $A(r) = \{x \in \mathbb{R}^{n-1} : (x, r) \in A\}$. We have

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

If $A = K^{\circ}$ then clearly A = -A and

$$A(-r) = \{x : (x, -r) \in A\} = \{x : (-x, r) \in A\} = \{-y : (y, r) \in A\} = -A(r).$$

In particular, $\operatorname{vol}_{n-1}(A(r)) = \operatorname{vol}_{n-1}(A(-r))$. By BM we get

$$\operatorname{vol}_{n-1}\left(\frac{K^{\circ}(r) + K^{\circ}(-r)}{2}\right) \ge \operatorname{vol}_{n-1}(K^{\circ}(r))^{1/2} \operatorname{vol}_{n-1}(K^{\circ}(-r))^{1/2} = \operatorname{vol}_{n-1}(K^{\circ}(r)).$$

We arrive at

$$\operatorname{vol}_{n}(S_{u}(K)^{\circ}) = \int_{-\infty}^{+\infty} \operatorname{vol}_{n-1}(S_{u}(K)^{\circ}(r)) dr \ge \int_{-\infty}^{+\infty} \operatorname{vol}_{n-1}\left(\frac{K^{\circ}(r) + K^{\circ}(-r)}{2}\right) dr$$
$$\ge \int_{-\infty}^{+\infty} \operatorname{vol}_{n-1}(K^{\circ}(r)) dr = \operatorname{vol}_{n}(K^{\circ}).$$

Rogers-Sheppard inequality. Note that the BM inequality implies

$$|K - K| \ge (|K|^{1/n} + |-K|^{1/n})^n = 2^n |K|.$$

Equality holds for convex bodies having centre of symmetry. We shall prove the reverse bound.

Theorem 9. Let K be a convex body in \mathbb{R}^n . Then

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

We need the following lemma.

Lemma 3. Let K, L be convex bodies in \mathbb{R}^n . Then the function

$$f(x) = \operatorname{vol}_n((x+L) \cap K)^{1/n}$$

is concave on its support K - L.

Proof. We check that K - L is indeed the support of f. If $(x + L) \cap K \neq \emptyset$ then are $l \in L$ and $k \in K$ such that x + l = k. Thus, $x = k - l \in K - L$.

It is straightforward to check that

$$\lambda((x+L)\cap K) + (1-\lambda)((y+L)\cap K) \subseteq (\lambda x + (1-\lambda)y + L) \cap K.$$

Thus, by Brunn-Minkowski

$$\begin{split} f(\lambda x + (1-\lambda)y) &= |\lambda((x+L)\cap K) + (1-\lambda)((y+L)\cap K)|^{1/n} \\ &\geq |\lambda((x+L)\cap K) + (1-\lambda)((y+L)\cap K)|^{1/n} \\ &\geq \lambda |(x+L)\cap K)|^{1/n} + (1-\lambda)|(y+L)\cap K)|^{1/n} \\ &= \lambda f(x) + (1-\lambda)f(y). \end{split}$$

- 1				L
	-	-	-	

Proof of the Rogers-Sheppard inequality. For $x \in K - K$ take the radial function

$$\rho_{K-K}(\theta) = \max\{t > 0: t\theta \in K - K\}, \qquad \theta \in S^{n-1}.$$

Take

$$f(x) = |K \cap (x+K)|^{1/n}, \qquad g(x) = f(0)\left(1 - \frac{r}{\rho_{K-K}(\theta)}\right), \quad x = r\theta.$$

We have f(0) = g(0). Moreover, both f and g vanish on the boundary of K - K (that is, when $r = \rho_{K-K}(\theta)$). Thus, since f is concave and g is linear in the radial coordinate, we get $f \ge g$ on K - K. Let $\kappa_n = \operatorname{vol}_n(B_2^n)| =$. Then $\operatorname{vol}_{n-1}(S^{n-1}) = n\kappa_n$. We get

$$\begin{split} \int_{K-K} |K \cap (x+K)| \mathrm{d}x &= \int_{K-K} f^n \mathrm{d}x \ge \int_{K-K} g^n \mathrm{d}x \\ &= f(0)^n n \kappa_n \int_{S^{n-1}} \int_0^{\rho_{K-K}(\theta)} r^{n-1} \left(1 - \frac{r}{\rho_{K-K}(\theta)}\right)^n \mathrm{d}r \mathrm{d}\sigma_{n-1}(\theta) \\ &= |K| n \kappa_n \int_{S^{n-1}} \rho_{K-K}(\theta)^n \mathrm{d}\sigma_{n-1}(\theta) \int_0^1 t^{n-1} (1-t)^n \mathrm{d}t \\ &= |K| \cdot |K - K| \frac{n \Gamma(n) \Gamma(n+1)}{\Gamma(2n+1)} = \binom{2n}{n}^{-1} |K| \cdot |K - K|. \end{split}$$

Here we have used the fact that

$$|K| = \kappa_n \int_{S^{n-1}} \rho_K^n(\theta) \mathrm{d}\sigma_{n-1}(\theta).$$

which follows from the fact that for a spherical cone we have

$$\operatorname{vol}_n(\mathrm{d}\theta) = \kappa_n r^n \mathrm{d}\sigma_{n-1}(\theta)$$

On the other hand we have

$$\int_{K-K} |K \cap (x+K)| \mathrm{d}x = \int_{\mathbb{R}^n} |K \cap (x+K)| \mathrm{d}x = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \mathbf{1}_K(y) \mathbf{1}_{K+x}(y) \mathrm{d}y \mathrm{d}x$$
$$= |K| \int_{\mathbb{R}^n} \mathbf{1}_K(y) \mathrm{d}y = |K|^2.$$

1.7. Spherical isoperimetry. Let $x_0 \in S^n$ and let H be a n-dimensional hyperplane in \mathbb{R}^{n+1} , not passing through x_0 . Then $\mathbb{R}^{n+1} \setminus H$ is a sum of two open halfspaces: H_+ containing x_0 and H_- not containing x_0 . Let i_H be the reflection through H. For a measurable $f: S^n \to I$ \mathbb{R} we define:

$$f^{H}(x) = \begin{cases} \max\{f(x), f(i_{H}x)\} & x \in H_{+} \\ \min\{f(x), f(i_{H}x)\} & x \in H_{-} \\ f(x) & x \in H \end{cases}$$

Let σ_n be the uniform measure on S^n . Let $\operatorname{dist}_{\sigma_n}(f)$ be the distribution of f under σ_n . We have the following lemma.

Lemma 4. We have

- (i) dist_{σ_n}(f^H) = dist_{σ_n}(f),
- (ii) if f is L-Lipschitz then f^H is also L-Lipschitz,
- (iii) $\int_{S^n} d(x, x_0) f(x) d\sigma_n(x) \ge \int_{S^n} d(x, x_0) f^H(x) d\sigma_n(x)$; moreover if f is continuous the equality holds if and only if $f = f^H$.

Proof. The first part is obvious. To prove the second part we first observe that since the minimum and maximum of two L-Lipschitz functions is L-Lipschitz, the function f^H is L-Lipschitz of H_+ and on H_- . It suffices to show that a continuous function f which is L-Lipschitz on H_+ and H_- is also L-Lipschitz on S^n . Suppose $x \in H_+$ and $y \in H_-$. Take $z \in H$ lying in the shortest geodesic between x and y. We have

$$|f(x) - f(y)| \le |f(x) - f(z)| + |f(z) - f(y)| \le L(|x - z| + |z - y|) = L|x - y|.$$

To prove (iii) we observe that

$$f(x) + f(i_H x) = f^H(x) + f^H(i_H x)$$

and thus

$$f(i_H x) - f^H(i_H x) = f^H(x) - f(x)$$

Thus

$$\begin{split} \int_{S^n} d(x_0, x) (f(x) - f^H(x)) \mathrm{d}\sigma_n(x) &= \int_{S^n \cap H^+} d(x_0, x) (f(x) - f^H(x)) \mathrm{d}\sigma_n(x) \\ &+ \int_{S^n \cap H^-} d(x_0, x) (f(x) - f^H(x)) \mathrm{d}\sigma_n(x) \\ &= \int_{S^n \cap H^+} d(x_0, x) (f(x) - f^H(x)) \mathrm{d}\sigma_n(x) + \int_{S^n \cap H^+} d(x_0, i_H x) (f(i_H x) - f^H(i_H x)) \mathrm{d}\sigma_n(x) \\ &= \int_{S^n \cap H^+} (d(x_0, x) - d(x_0, i_H x)) (f(x) - f^H(x)) \mathrm{d}\sigma_n(x) \ge 0, \end{split}$$
where on H^+ we have $f(x) \leq f^H(x)$ and $d(x_0, x) \leq d(x_0, i_H x).$

since on H^+ we have $f(x) \leq f^H(x)$ and $d(x_0, x) \leq d(x_0, i_H x)$.

Lemma 5. The radial symmetrization f^* always exists. Moreover, $f^* = f$ if and only if $f^H = f$ for any H.

Proof. Let $F(t) = \sigma_n (f \leq t)$ be the distribution function of f. It is a standard exercise to show that the function

$$f_0(s) = \sup\{u : F(u) < s\}, \quad s \in (0, 1)$$

defines a non-decreasing function which, viewed as a function $f_0: (0,1) \to \mathbb{R}$ defined on (0,1)with Lebesgue measure, defines a random variable whose distribution function is equal to F. Let T be a radial map pushing forward σ_n onto Leb_[0,1]. This map is defined via

$$x \to \sigma_n(\{y : d(y, x_0) \le d(x, x_0)\}).$$

The function $f^* = f_0 \circ (1 - T)$ is the desired radial symmetrization. Indeed f^* is radially non-increasing and 1 - T also pushes forward σ_n onto $\text{Leb}_{[0,1]}$. Thus

$$\sigma_n(f^* \le t) = |\{f_0 \le t\}| = F(t) = \sigma_n(f \le t).$$

To prove the second claim we first observe that if f is radial, then clearly $f = f^H$ for any H. To prove the other implication assume that f is not radial. Then there are point $x, y \in S^n$ such that $d(x_0, x) \leq d(x_0, y)$ and f(x) < f(y). Take the segment [x, y] and bisect it with the hyperplane H. Clearly $x \in H_+$ and $y \in H_-$. We have

$$f^{H}(x) = \max(f(x), f(i_{H}x)) = \max(f(x), f(y)) = f(y) > f(x).$$

Thus $f^H \neq f$.

We need yet another lemma.

Lemma 6. Let us fix an *L*-Lipschitz function $f: S^n \to \mathbb{R}$. Define

$$\mathcal{A} = \{g: S^n \to \mathbb{R} : \text{ dist}_{\sigma_n}(g) = \text{dist}_{\sigma_n}(f), g \text{ is } L - \text{Lipschitz}\}$$

Take $m = \inf_{g \in \mathcal{A}} \int_{S^n} d(x_0, x) g(x) d\sigma_n(x)$. Then

(i) There is a sequence $(g_k) \subset \mathcal{A}$ and a function $g \in \mathcal{A}$ such that $g_k \to g$ uniformly and

$$\lim_{k \to \infty} \int_{S^n} d(x_0, x) g_k(x) \mathrm{d}\sigma_n(x) = m.$$

- (ii) We have $\int_{S^n} d(x_0, x) g(x) d\sigma_n(x) = m$.
- (iii) We have $g = f^*$.

Proof. Clearly there is a sequence (g_k) such that

$$\lim_{k \to \infty} \int_{S^n} d(x_0, x) g_k(x) \mathrm{d}\sigma_n(x) = m.$$

Part (i) follows from Arzela-Ascoli theorem (equicontinuity follows from the fact that the members of \mathcal{A} are *L*-Lipschitz and pointwise boundedness from the fact that for any $g \in \mathcal{A}$ we have $\operatorname{dist}_{\sigma_n}(g) = \operatorname{dist}_{\sigma_n} f$ and f is bounded). Clearly g is *L*-Lipschitz as a pointwise limit of *L*-Lipschitz functions. Let $\mathbf{1}_{g_n < t}(x) \to \mathbf{1}_{g < t}(x)$ for any x we get by the Lebesgue dominated convergence theorem (and Lemma 4 (i)) that $\sigma_n(f < t) = \sigma_n(g_n < t) \to \sigma_n(g < t)$, which implies $\operatorname{dist}_{\sigma_n}(g) = \operatorname{dist}_{\sigma_n}(f)$.

Part (ii) follows by the Lebesgue dominated convergence theorem.

To prove part (iii), in view of Lemma 5, it suffices to prove that $g = g^H$ for any H. If $g \neq g^H$ for some H then by continuity of g we get (Lemma 4 (iii))

$$m = \int_{S^n} d(x_0, x) g(x) \mathrm{d}\sigma_n(x) > \int_{S^n} d(x_0, x) g^H(x) \mathrm{d}\sigma_n(x)$$

This contradicts the minimality of m.

We are ready to prove the spherical isoperimetric inequality. A set $C \subseteq S^n$ of the form $C_{x_0,t} = \{x \in S^n : d(x_0, x) \leq t\}$ is called a spherical cup.

Theorem 10. Suppose A is a compact set. Let C be a spherical cup such that $\sigma_n(A) = \sigma_n(C)$. Then $\sigma_n(A_t) \ge \sigma_n(C_t)$.

Proof. Define $f(x) = \max\{t - d(x, A), 0\}$. The set $\{t \ge f > 0\}$ is the open t-enlargement of A. We have

$$\sigma_n(A_t) = \sigma_n(\{t \ge f > 0\}) = \sigma_n(\{t \ge f^* > 0\})$$

Let A^* be the spherical cup centred at x_0 given by $A^* = \{f^* = t\}$. We have

$$\sigma_n(A^*) = \sigma_n(f^* = t) = \sigma_n(f = t) = \sigma_n(A).$$

If $x \in A^*$ then $f^*(x) = t$ and since by Lemma 6 the function f^* is 1-Lipschitz (as f was 1-Lipschitz) we get that $d(x, A^*) < t$ implies $f^*(x) > 0$. We arrive at

$$\sigma_n(A_t) = \sigma_n(\{t \ge f > 0\}) = \sigma_n(\{t \ge f^* > 0\}) \ge \sigma_n((A^*)_t).$$

We shall deduce the Gaussian isoperimetric inequality from the spherical isoperimetry.

Theorem 11. Suppose A is a compact set in \mathbb{R}^k . Let H be a half-space satisfying $\gamma_k(A) = \gamma_k(H)$. Then $\gamma_k(A_t) \geq \gamma_k(H_t)$.

We need the following well-known lemma.

Lemma 7. Let $\tilde{\sigma}_{n-1}$ be the uniform measure on $\sqrt{n}S^{n-1}$. Let $\pi_{k,n}(x_1,\ldots,x_n) = (x_1,\ldots,x_k)$ be the standard projection $\mathbb{R}^n \to \mathbb{R}^k$. Define $\mu_{k,n} = \pi_{k,n}\tilde{\sigma}_{n-1}$, that is

$$\mu_{k,n}(A) = \tilde{\sigma}_{n-1}(\pi_{k,n}^{-1}(A))$$

Then $\mu_{k,n} \to \gamma_k$ in the sense of distribution.

Proof. Let g_1, g_2, \ldots be independent standard normal real random variables. Since

$$\sigma_{n-1} \sim \frac{(g_1, \dots, g_n)}{|(g_1, \dots, g_n)|} = \frac{(g_1, \dots, g_n)}{(g_1^2 + \dots + g_n^2)^{1/2}},$$

we have

$$\pi_{k,n}\tilde{\sigma}_{n-1}\sim\sqrt{n}\frac{(g_1,\ldots,g_k)}{(g_1^2+\ldots+g_n^2)^{1/2}}=\alpha_n(g_1,\ldots,g_k),$$

where

$$\alpha_n = \left(\frac{g_1^2 + \ldots + g_n^2}{n}\right)^{1/2} \to 1 \qquad a.s$$

The assertion follows.

Sketch of the proof of Gaussian isoperimetry. We can assume that H is of the form $H = \{x_1 \ge r\}$. Since $\pi_{k,n}$ is 1-Lipschitz, it is easy to verify that $(\pi_{k,n}^{-1}(A))_t \subseteq \pi_{k,n}^{-1}(A_t)$. Thus

$$\gamma_k(A) \ge \limsup_{n \to \infty} \mu_{k,n}(A) = \limsup_{n \to \infty} \tilde{\sigma}_{n-1}(\pi_{k,n}^{-1}(A_t)) \ge \limsup_{n \to \infty} \tilde{\sigma}_{n-1}((\pi_{k,n}^{-1}(A))_t).$$

Since $\tilde{\sigma}_{n-1}(\pi_{k,n}^{-1}(A)) = \tilde{\sigma}_{n-1}(\pi_{k,n}^{-1}(H))$ we get from the spherical isoperimetry that

$$\limsup_{n \to \infty} \tilde{\sigma}_{n-1}((\pi_{k,n}^{-1}(A))_t) \ge \limsup_{n \to \infty} \tilde{\sigma}_{n-1}((\pi_{k,n}^{-1}(H))_t) = \limsup_{n \to \infty} \mu_{k,n}(H_{t+\varepsilon_n}) = \gamma_k(H_t),$$

where $\varepsilon_n \to 0$ is some explicitly computable sequence.

Corollary 12. Suppose A is a compact set in \mathbb{R}^n . Let $H = \{x_1 \leq r\}$ be such that $\gamma_n(A) = \gamma_n(H) = \Phi(r)$, where $\Phi(r) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^r e^{-x^2/2} dx$. Then

$$\gamma_n(A_t) \ge \Phi(r_t) = \Phi\left(\Phi^{-1}(\gamma_n(A)) + t\right).$$

In particular,

$$\gamma_n(A) \ge 1/2 \qquad \Longrightarrow \qquad \gamma_n(A_t) \ge \Phi(t).$$

Taking the derivative in t at t = 0 we get

$$\gamma_n(\partial A) \ge \Phi'\left(\Phi^{-1}(\gamma_n(A)) + t\right) = \varphi\left(\Phi^{-1}(\gamma_n(A))\right) = I(\gamma_n(A)),$$

where $I = \varphi \circ \Phi^{-1}$ and $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

Theorem 13 (Bobkov's inequality). Suppose $f : \mathbb{R}^n \to [0, 1]$ is a smooth function. Then

$$I\left(\int f \mathrm{d}\gamma_n\right) \leq \int \sqrt{I(f)^2 + |\nabla f|^2} \mathrm{d}\gamma_n.$$

Proof. Let

$$A = \{(x, y): x \in \mathbb{R}^n, y \in \mathbb{R}, \Phi(y) < f(x)\}.$$

Let $g = \Phi^{-1} \circ f$. Then

$$\gamma_{n+1}(A) = \int_{\mathbb{R}^n} \int_{-\infty}^{g(x)} \mathrm{d}\gamma_1(y) \mathrm{d}\gamma_n(x) = \int_{\mathbb{R}^n} \Phi(g(x)) \mathrm{d}\gamma_n(x) = \int_{\mathbb{R}^n} f \mathrm{d}\gamma_n(x) \mathrm$$

Moreover, if $\varphi_n(x) = (2\pi)^{-n/2} e^{-|x|^2/2}$ then

$$\gamma_{n+1}(\partial A) = \int_{\mathbb{R}^n} \varphi_n(x)\varphi(g(x))\sqrt{1+|\nabla g|^2} dx = \int \sqrt{\varphi(g)^2+\varphi(g)^2}|\nabla g|^2$$
$$= \int \sqrt{\varphi(g)^2+|\nabla\Phi\circ g|^2} = \int \sqrt{I(f)^2+|\nabla f|^2}.$$

Thus the assertion follows from the Gaussian isoperimetric inequality $\gamma_{n+1}(\partial A) \ge I(\gamma_{n+1}(A))$.

 \square

1.8. **Prékopa-Leindler inequality.** We are going to prove the following fundamental theorem.

Theorem 14. (PREKOPA-LEINDLER, '88) Let f, g, m be nonnegative measurable functions on \mathbb{R}^n and let $\lambda \in [0, 1]$. If for all $x, y \in \mathbb{R}^n$ we have

$$m((1-\lambda)x + \lambda y) \ge f(x)^{1-\lambda}g(y)^{\lambda},$$

then

(6)
$$\int_{\mathbb{R}^n} m \ge \left(\int_{\mathbb{R}^n} f\right)^{1-\lambda} \left(\int_{\mathbb{R}^n} g\right)^{\lambda}.$$

We first give two proof of this fact in dimension n = 1.

First proof of Prékopa-Leinlder in dimension one. We start with proving (B-M) inequality in dimension 1. Let A, B be compact sets in \mathbb{R} . Observe that the operations $A \to A + v_1$, $B \to B + v_2$ where $v_1, v_2 \in \mathbb{R}$ does not change the volumes of A, B and $(1 - \lambda)A + \lambda B$ (adding a number to one of the sets only shifts all of this sets). Therefore we can assume that $\sup A = \inf B = 0$. But then, since $0 \in A$ and $0 \in B$, we have

$$(1 - \lambda)A + \lambda B \supset (1 - \lambda)A \cup (\lambda B).$$

But $(1 - \lambda)A$ and (λB) are disjoint, up to the one point 0. Therefore

$$|(1-\lambda)A + \lambda B| \ge |(1-\lambda)A| + |\lambda B|,$$

hence we have proved (B-M) in dimension 1.

Let us now justify the Prekopa-Leindler inequality in dimension 1. We can assume, considering $f\mathbf{1}_{f\leq M}$ and $g\mathbf{1}_{g\leq M}$ instead of f and g, that f, g are bounded. Note also that this inequality possesses some homogeneity. Indeed, if we multiply f, g, m by numbers c_f, c_g, c_m satisfying

$$c_m = c_f^{1-\lambda} c_g^\lambda,$$

then the hyphotesis and the thesis do not change. Therefore, taking $c_f = ||f||_{\infty}^{-1}$, $c_g = ||g||_{\infty}^{-1}$ and $c_m = ||f||_{\infty}^{-(1-\lambda)} ||g||_{\infty}^{-\lambda}$ we can assume (since we are in the situation when f and g are bounded) that $||f||_{\infty} = ||g||_{\infty} = 1$. But then

$$\int_{\mathbb{R}} m = \int_{0}^{+\infty} |\{m \ge s\}| \, \mathrm{d}s,$$
$$\int_{\mathbb{R}} f = \int_{0}^{1} |\{f \ge r\}| \, \mathrm{d}r,$$
$$\int_{\mathbb{R}} g = \int_{0}^{1} |\{g \ge r\}| \, \mathrm{d}r.$$

Note also that if $x \in \{f \ge r\}$ and $y \in \{g \ge r\}$ then by the assumption of the theorem we have $(1 - \lambda)x + \lambda y \in \{m \ge r\}$. Hence,

$$(1-\lambda)\{f \ge r\} + \lambda\{g \ge r\} \subset \{m \ge r\}.$$

Moreover, the sets $\{f \ge r\}$ and $\{g \ge r\}$ are non-empty for $r \in [0, 1)$. This is very important since we want to use 1 dimensional (B-M) inequality! We have

$$\int m = \int_0^{+\infty} |\{m \ge r\}| \, \mathrm{d}r \ge \int_0^1 |\{m \ge r\}| \, \mathrm{d}r \ge \int_0^1 |(1-\lambda)\{f \ge r\} + \lambda\{g \ge r\}| \, \mathrm{d}r$$
$$\ge (1-\lambda) \int_0^1 |\{f \ge r\}| \, \mathrm{d}r + \lambda \int_0^1 |\{g \ge r\}| \, \mathrm{d}r = (1-\lambda) \int f + \lambda \int g$$
$$\ge \left(\int f\right)^{1-\lambda} \left(\int g\right)^{\lambda}.$$

Observe that we have proved

$$\int m \ge (1-\lambda) \int f + \lambda \int g,$$

but this inequality does not have the previous homogeneity, hence it requires the assumption $||f||_{\infty} = ||g||_{\infty} = 1.$

For the second proof we shall assume that f, g, m are strictly positive and smooth.

Second proof of Prékopa-Leinlder in dimension one. Assume without loss of generality that $\int f = F > 0$ and $\int g = G > 0$. Define $x, y : [0, 1] \to \mathbb{R}$ such that x(t), y(t) are the infima of numbers satisfying

$$\frac{1}{F}\int_{-\infty}^{x(t)} f(s)\mathrm{d}s = \frac{1}{G}\int_{-\infty}^{y(t)} g(s)\mathrm{d}s = t.$$

The functions x, y are differentiable due to our assumptions. Define $z(t) = \lambda x(t) + (1 - \lambda)y(t)$. Differentiating the above equalities we get

$$\frac{f(x(t))x'(t)}{F} = \frac{g(y(t))y'(t)}{G} = 1.$$

Thus, using the assumption of Prékopa-Leindler together with AM-GM we get

$$\begin{split} \int h &\geq \int_0^1 h(z(t)) z'(t) \mathrm{d}t \geq \int_0^1 h(\lambda x(t) + (1-\lambda)y(t)) (\lambda x'(t) + (1-\lambda)y'(t)) \\ &\geq \int_0^1 f(x(t))^\lambda g(y(t))^{1-\lambda} x'(t)^\lambda y'(t)^{1-\lambda} \\ &= \int_0^1 (f(x(t))x'(t))^\lambda (g(y(t))y'(t))^{1-\lambda} \\ &= F^\lambda G^{1-\lambda} = \left(\int f\right)^\lambda \left(\int g\right)^{1-\lambda}. \end{split}$$

Proof of Prékopa-Leindler in dimension n > 1. Suppose our inequality in true in dimension n-1. We will prove it in dimension n. Suppose we have a numbers $y_0, y_1, y_2 \in \mathbb{R}$ satisfying $y_0 = (1-\lambda)y_1 + \lambda y_2$. Define $m_{y_0}, f_{y_1}, g_{y_2} : \mathbb{R}^{n-1} \to \mathbb{R}_+$ by

$$m_{y_0}(x) = m(y_0, x), \quad f_{y_1}(x) = f(y_1, x), \quad g_{y_2}(x) = (y_2, x),$$

where $x \in \mathbb{R}^{n-1}$. Note that since $y_0 = (1 - \lambda)y_1 + \lambda y_2$ we have

$$m_{y_0}((1-\lambda)x_1 + \lambda x_2) = m((1-\lambda)y_1 + \lambda y_2, (1-\lambda)x_1 + \lambda x_2)$$

$$\geq f(y_1, x_1)^{1-\lambda}g(y_2, x_2)^{\lambda} = f_{y_1}(x_1)^{1-\lambda}g_{y_2}(x_2)^{\lambda},$$

hence m_{y_0}, f_{y_1} and g_{y_2} satisfies the assumption of the (n-1)-dimensional Prekopa-Leindler inequality. Therefore we have

$$\int_{\mathbb{R}^{n-1}} m_{y_0} \ge \left(\int_{\mathbb{R}^{n-1}} f_{y_1}\right)^{1-\lambda} \left(\int_{\mathbb{R}^{n-1}} g_{y_2}\right)^{\lambda}.$$

Step 4. Define new functions $M, F, G : \mathbb{R} \to \mathbb{R}_+$

$$M(y_0) = \int_{\mathbb{R}^{n-1}} m_{y_0}, \quad F(y_1) = \int_{\mathbb{R}^{n-1}} f_{y_1}, \quad G(y_2) = \int_{\mathbb{R}^{n-1}} g_{y_2}.$$

We have seen (the above inequality) that when $y_0 = (1 - \lambda)y_1 + \lambda y_2$ then there holds

$$M((1-\lambda)y_1+\lambda y_2) \ge F(y_1)^{1-\lambda}G(y_2)^{\lambda}.$$

Hence, by 1-dimensional (P-L) inequality we get

$$\int_{\mathbb{R}} M \ge \left(\int_{\mathbb{R}} F\right)^{1-\lambda} \left(\int_{\mathbb{R}} G\right)^{\lambda}.$$

But

$$\int_{\mathbb{R}} M = \int_{\mathbb{R}^n} m, \quad \int_{\mathbb{R}} F = \int_{\mathbb{R}^n} f, \quad \int_{\mathbb{R}} G = \int_{\mathbb{R}^n} g,$$

so we shown that

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

Definition 1. A function $f : \mathbb{R}^n \to R$ is called log-concave if $f = e^{-V}$ for some convex function $V : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$.

We can now give a proof of generalization of BM inequality.

Theorem 15. Suppose μ is a measure with log-concave density. Then

$$\mu(\lambda A + (1 - \lambda)B) \ge \mu(A)^{\lambda}\mu(B)^{1 - \lambda}.$$

Proof. Let A, B be measurable in \mathbb{R}^n and let h be the density of μ . Define $f = \mathbf{1}_A h, g = \mathbf{1}_B h$ and $m = \mathbf{1}_{\lambda A + (1-\lambda)B} h$. Then these function clearly satisfy $m(\lambda x + (1-\lambda)y) \ge f(x)^{\lambda}g(y)^{\lambda}$. Thus

$$|\lambda A + (1-\lambda)B| = \int m \ge \left(\int f\right)^{\lambda} \left(\int g\right)^{1-\lambda} = |A|^{\lambda} |B|^{1-\lambda}.$$

Fact 16. Suppose $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is log-concave. Then $F(x) = \int_{\mathbb{R}^m} f(x, y) dy$ is also log-concave.

Proof. Define $f_x(y) = f(x, y), f_x : \mathbb{R}^m \to \mathbb{R}$. Take $x_1, x_2 \in \mathbb{R}^n$. The functions $f_{\lambda x_1 + (1-\lambda)x_2}, f_{x_1}, f_{x_2}$ satisfy

$$f_{\lambda x_1 + (1-\lambda)x_2}(\lambda y_1 + (1-\lambda)y_2) \ge f_{x_1}(y_1)^{\lambda} f_{x_2}(y_2)^{1-\lambda}$$

Thus by Prékopa-Leindler

$$F(\lambda x_1 + (1-\lambda)x_2) = \int f_{\lambda x_1 + (1-\lambda)x_2} \ge \left(\int f_{x_1}\right)^{\lambda} \left(\int f_{x_2}\right)^{1-\lambda} = F(x_1)^{\lambda} F(x_2)^{1-\lambda}.$$

Fact 17. Let f, g be log-concave on \mathbb{R}^n . Then f * g is also log-concave.

Proof. The function $(x, y) \to f(y)g(x-y)$ is clearly log concave. Thus it suffices to integrate it in y and use Fact 16.

Fact 18. Let f be log-concave on \mathbb{R}^n and let $v \in \mathbb{R}^n$ be a fixed vector.

$$\mathbb{R} \ni t \longmapsto \int_{\langle x, v \rangle \ge t} f(x) \mathrm{d}x$$

is also log-concave.

Proof. The function $(x,t) \mapsto f(x) \mathbf{1}_{\langle x,v \rangle \geq t}$ is log-concave (the function $(x,t) \mapsto \mathbf{1}_{\langle x,v \rangle \geq t}$ is log-concave as it is of the form $\mathbf{1}_K$ for a convex K with K being a half-space). It suffices to use Fact 16.

Gaussian concentration. We shall prove the following fact.

Theorem 19. Let $A \subset \mathbb{R}^n$ and let γ_n be the Gaussian measure. Then

(7)
$$\int \exp\left(\frac{d(x,A)^2}{4}\right) d\gamma_n(x) \le \frac{1}{\gamma_n(A)}.$$

Moreover, if $\gamma_n(A) \ge 1/2$ then

(8)
$$\gamma_n(A_{\varepsilon}) \ge 1 - 2\exp(-\varepsilon^2/4).$$

Proof. Let

$$f(x) = \frac{1}{(2\pi)^{n/2}} \exp(d(x, A)^2/4) \exp(-|x|^2/2),$$
$$g(y) = \frac{1}{(2\pi)^{n/2}} \mathbf{1}_A(y) \exp(-|y|^2/2)$$

and

$$h(z) = \frac{1}{(2\pi)^{n/2}} \exp(-|z|^2/2).$$

We show that

$$h\left(\frac{x+y}{2}\right) \ge \sqrt{f(x)}\sqrt{g(y)}.$$

Indeed, it suffices to consider the case when $y \in A$. In this case we have $d(x, A) \leq |x - y|_2$ and therefore

$$(2\pi)^n f(x)g(y) \le \exp\left(\frac{|x-y|^2}{4} - \frac{|x|^2}{2} - \frac{|y|^2}{2}\right) = \exp\left(-\frac{|x+y|^2}{4}\right)$$
$$= (2\pi)^n \left(h\left(\frac{x+y}{2}\right)\right)^2.$$

By the Prékopa-Leindler inequality we obtain

$$1 = \left(\int h\right)^2 \ge \left(\int f\right) \left(\int g\right) = \gamma_n(A) \int \exp\left(\frac{d(x,A)^2}{4}\right) d\gamma_n(x).$$

The second part of the statement follows from Markov's inequality. Indeed, if $\gamma_n(A) \ge 1/2$ then

$$\int \exp(d(x,A)^2/4) \mathrm{d}\gamma_n(x) \le 2$$

hence

$$\gamma_n(d(x,A) \ge \varepsilon) \le \exp(-\varepsilon^2/4) \int \exp\left(\frac{d(x,A)^2}{4}\right) d\gamma_n(x) \le 2\exp(-\varepsilon^2/4).$$

Corollary 20. If M is a γ_n median of a 1-Lipschitz function f, then

$$\gamma_n(\{f \ge M + \varepsilon\}) \le 2\exp(-\varepsilon^2/4), \ \gamma_n(\{f \le M - \varepsilon\}) \le 2\exp(-\varepsilon^2/4),$$

and

$$\gamma_n(\{|f-M| \ge \varepsilon\}) \le 4\exp(-\varepsilon^2/4).$$

Proof. Let $A = \{f \leq M\}$. Then $\gamma_n(A) \geq 1/2$. Since f is 1-Lipschitz we have $\{f \geq M + \varepsilon\} \subset A^c_{\varepsilon}$. Therefore,

$$\gamma_n(\{f \ge M + \varepsilon\}) \le \gamma_n(A_{\varepsilon}^c) \le 2\exp(-\varepsilon^2/4)$$

The second inequality is proven identically, taking $A = \{f \leq M\}$.

One can provide a nice estimate of the volume of a cup.

Fact 21. Let σ_{n-1} be the uniform probability measure on S^{n-1} . Take

$$C(\varepsilon) = S^{n-1} \cap \{x_1 \ge \varepsilon\}.$$

Then $\sigma_{n-1}(\varepsilon) \leq e^{-n\varepsilon^2/2}$.

Proof. Case 1. Assume $\varepsilon \in [0, 1/\sqrt{2}]$. Let $C = \operatorname{conv}(0, C(\varepsilon))$. Notice that and let $C \subseteq B(\varepsilon, \sqrt{1-\varepsilon^2})$. We thus have

$$\sigma_{n-1}(C(\varepsilon)) = \frac{\operatorname{vol}_n(C)}{\operatorname{vol}_n(B_2^n)} \le \frac{B(\varepsilon, \sqrt{1-\varepsilon^2})}{\operatorname{vol}_n(B_2^n)} = (1-\varepsilon^2)^{n/2} \le e^{-n\varepsilon^2/2}.$$

If $\varepsilon \in [1/\sqrt{2}, 1]$ then $C \subseteq B(\frac{1}{2\varepsilon}, \frac{1}{2\varepsilon})$ and thus

$$\sigma_{n-1}(C(\varepsilon)) = \frac{\operatorname{vol}_n(C)}{\operatorname{vol}_n(B_2^n)} \le \frac{B(\frac{1}{2\varepsilon}, \frac{1}{2\varepsilon})}{\operatorname{vol}_n(B_2^n)} = \left(\frac{1}{2\varepsilon}\right)^n \le e^{-n\varepsilon^2/2}.$$

The last inequality follows from $e^{x^2/2} < 2x$ for $x \in [1/\sqrt{2}, 1]$, which is easy to verify. In fact due to the convexity of $e^{x^2/2}$ it is enough to check it for $x = 1/\sqrt{2}$ and x = 1. In these two cases the inequality easily reduces to e < 4.

Fact 22. Suppose X is a random vector having values in \mathbb{R}^n , whose density is of the form

$$g(x) = \exp\left(-\frac{1}{2}\langle Bx, x\rangle - V(x)\right),$$

where $B \ge 0$ is a $n \times n$ matrix and V is convex. Then

$$\operatorname{cov}(X) \le B^{-1}$$

19

In other words, of X is more log concave than a Gaussian vector Y then $cov(X) \leq cov(Y)$.

Proof. Writing $X = B^{-1/2} \tilde{X}$ we clearly see that one can assume the case B = I. Indeed, \tilde{X} has density

$$\tilde{g}(x) = \det(B^{-1/2})g(B^{-1/2}x) = \det(B^{-1/2})\exp\left(-\frac{1}{2}|x|^2 - V(B^{-1/2}x)\right)$$

and $\operatorname{cov}(\tilde{X}) = B \operatorname{cov}(X) = BB^{-1} = I.$ Let $\Lambda(y) = \log \mathbb{E}e^{\langle y, X \rangle}$. We have

$$\frac{\partial^2}{\partial y_i y_j} \Lambda(y) = \frac{\mathbb{E} X_i X_j e^{\langle y, X \rangle} \mathbb{E} e^{\langle y, X \rangle} - \mathbb{E} X_i e^{\langle y, X \rangle} \mathbb{E} X_j e^{\langle y, X \rangle}}{(\mathbb{E} e^{\langle x, X \rangle})^2}$$

Thus $\frac{\partial^2}{\partial y_i y_j} \Lambda(y) \Big|_{y=0} = \operatorname{cov}(X_i, X_j)$. Thus $\nabla^2 \Lambda(0) = \operatorname{cov}(X)$. Let us define

$$f(x) = \langle a, x \rangle - \frac{1}{2} |x|^2 - V(x)$$
$$g(y) = -\langle a, y \rangle - \frac{1}{2} |y|^2 - V(y)$$
$$m(z) = -\frac{1}{2} - V(z).$$

We shall verify the inequality

$$\frac{1}{2}f(x) + \frac{1}{2}g(y) \le \frac{1}{2}|a|^2 + m\left(\frac{x+y}{2}\right).$$

Indeed, due to convexity of V it is enough to check

$$\frac{1}{2} \left\langle a, x \right\rangle - \frac{1}{4} |x|^2 - \frac{1}{2} \left\langle a, y \right\rangle - \frac{1}{4} |y|^2 \le \frac{1}{2} |a|^2 - \frac{1}{2} \left| \frac{x+y}{2} \right|^2.$$

This is

$$\begin{split} \left\langle a, \frac{x-y}{2} \right\rangle &\leq \frac{1}{4} |x|^2 + \frac{1}{4} |y|^2 + \frac{1}{2} |a|^2 - \frac{1}{8} |x+y|^2 = \frac{1}{2} |a|^2 + \frac{1}{8} |x|^2 + \frac{1}{2} |y|^2 - \frac{1}{4} \left\langle x, y \right\rangle \\ &= \frac{1}{2} |a|^2 + \frac{1}{2} \left| \frac{x-y}{2} \right|^2. \end{split}$$

The inequality follow by applying Cauchy-Schwarz and AM-GM,

$$\left\langle a, \frac{x-y}{2} \right\rangle \le |a| \left| \frac{x-y}{2} \right| \le \frac{1}{2} |a|^2 + \frac{1}{2} \left| \frac{x-y}{2} \right|^2.$$

Thus

$$\sqrt{e^{f(x)}e^{g(y)}} \le e^{\frac{1}{2}|a|^2}e^{m\left(\frac{x+y}{2}\right)}.$$

Thus by Prékopa-Leindler we get

$$\left(\int e^f\right)^{1/2} \left(\int e^g\right)^{1/2} \le e^{\frac{1}{2}|a|^2} \int e^m,$$

which is equivalent to

$$\frac{1}{2}\Lambda(a) + \frac{1}{2}\Lambda(-a) - \Lambda(0) \le \frac{1}{2}|a|^2.$$

If we Taylor expand the left hand side we get

$$\left<\nabla^2\Lambda(0)a,a\right>+o(|a|^2)\leq |a|^2$$

and after comparing the leading terms we get $\langle \nabla^2 \Lambda(0)a, a \rangle \leq |a|^2$ which shows that $I - \Lambda(0) \geq 0$.

1.9. Knothe map.

1.10. Brenier map.

1.11. Ehrhard inequality. Recall that γ_n is the standard Gaussian measure on \mathbb{R}^n and $\Phi(t) = \int \varphi(s) ds$, there φ is the density of γ_1 . The main goal of this section is to prove the following theorem, known as Ehrhard inequality.

Theorem 23. Suppose A, B are Borel sets in \mathbb{R}^n . Suppose $\alpha, \beta > 0$ are such that $\alpha + \beta \ge 1$ and $|\alpha - \beta| \le 1$. Then we have

$$\Phi^{-1}(\gamma_n(\alpha A + \beta B)) \ge \alpha \Phi^{-1}(\gamma_n(A)) + \beta \Phi^{-1}(\gamma_n(B)).$$

Let us introduce the functional form of this inequality. Define the operator

$$(Q_t f)(x) = \int_{\mathbb{R}^n} f(x + \sqrt{t}z) \mathrm{d}\gamma_n(z).$$

Note that

$$(Q_1 f)(0) = \int f \mathrm{d}\gamma_n$$

Moreover, let us observe that $f \ge c$ implies $Q_t f \ge c$ and $f \le C$ implies $Q_t f \le C$.

Theorem 24. Suppose $\alpha, \beta > 0$ are such that $\alpha + \beta \ge 1$ and $|\alpha - \beta| \le 1$. Let $f, g, h : \mathbb{R}^n \to (0, 1)$ be Borel functions such that

$$\Phi^{-1}(h(\alpha x + \beta y)) \ge \alpha \Phi^{-1}(f(x)) + \beta \Phi^{-1}(g(y)).$$

Then

(9)
$$\Phi^{-1}(Q_t h(\alpha x + \beta y)) \ge \alpha \Phi^{-1}(Q_t f(x)) + \beta \Phi^{-1}(Q_t g(y)).$$

In particular, taking x = y = 0 and t = 1 yields

$$\Phi^{-1}\left(\int h \mathrm{d}\gamma_n\right) \ge \alpha \Phi^{-1}\left(\int f \mathrm{d}\gamma_n\right) + \beta \Phi^{-1}\left(\int g \mathrm{d}\gamma_n\right).$$

It is not hard to show that this theorem implies Ehrhard inequality. However, we shall not need this implication. We first show that Erhard inequality implies Theorem 24.

Theorem 23 implies 24. Consider a Borel set

$$B_f^x = \{(s,z) \in \mathbb{R} \times \mathbb{R}^n : s \le \Phi^{-1}(f(x+\sqrt{t}z))\}$$

Similarly we define B_g^y and $B_h^{\alpha x + \beta y}$. Observe that

$$\gamma_{n+1}(B_f^x) = \int f(x + \sqrt{t}z) \mathrm{d}\gamma_n(z) = Q_t f(z).$$

Thus, our goal (9) is equivalent to

$$\Phi^{-1}(\gamma_{n+1}(B_h^{\alpha x+\beta y})) \ge \alpha \Phi^{-1}(\gamma_{n+1}(B_f^x)) + \beta \Phi^{-1}(\gamma_{n+1}(B_g^y)).$$

But Ehrhard inequality gives

$$\Phi^{-1}(\gamma_{n+1}(\alpha B_f^x + \beta B_g^y)) \ge \alpha \Phi^{-1}(\gamma_{n+1}(B_f^x)) + \beta \Phi^{-1}(\gamma_{n+1}(B_g^y)).$$

Thus, it is enough to verify

$$\alpha B_f^x + \beta B_g^y \subseteq B_h^{\alpha x + \beta y}.$$

If $(s_1, z_1) \in B_f^x$ and $(s_2, z_2) \in B_g^y$ then from our assumptions we get $\alpha s_1 + \beta s_2 \leq \alpha \Phi^{-1}(f(x + \sqrt{t}z_1)) + \beta \Phi^{-1}(g(y + \sqrt{t}z_2)) \leq \Phi^{-1}(h(\alpha x + \beta y + \sqrt{t}(\alpha z_1 + \beta z_2)))),$ which show that $\alpha(s_1, z_1) + \beta(s_2, z_2) = (\alpha s_1 + \beta s_2, \alpha z_1 + \beta z_2) \in B_h^{\alpha x + \beta y}.$

Now our goal is to show that Theorem 24 for *nice functions* implies Ehrhard inequality. Let us first specify that we mean by nice functions. Take parameters a > 0, $0 < 2\varepsilon < \rho < 1$. Define

$$\delta_{\varepsilon,\rho} = \max\left\{\Phi(\alpha\Phi^{-1}(2\varepsilon) + \beta\Phi^{-1}(\rho)), \Phi(\alpha\Phi^{-1}(\rho) + \beta\Phi^{-1}(2\varepsilon))\right\}$$

and

$$\mathcal{N}_{a,\varepsilon,\rho} = \left\{ (f,g,h) : f,g,h : \mathbb{R}^n \to (0,1) \text{ are } C^\infty \text{ smooth} \\ f,g = \varepsilon \text{ outside } B(0,a) \\ f,g \le \rho \text{ everywhere} \\ h \ge \delta_{\varepsilon,\rho} \text{ everywhere} \right\}.$$

Suppose we know Theorem 24 for these classes of functions. We now show how to deduce Ehrhard inequality.

Theorem 24 for $\mathcal{N}_{a,\varepsilon,\rho}$ implies Theorem 23. By a standard reasoning similar to that discussed in the context of classical Brunn-Minkowski shows that we can restrict ourselves to compact sets. Take $0 < 2\varepsilon < \rho < 1$ and some $\eta > 0$. There are smooth functions f, g, h such that

$$f = \begin{cases} \rho & \text{on } A \\ \varepsilon & \text{on } A_{\eta}^{c} \end{cases}, \qquad g = \begin{cases} \rho & \text{on } B \\ \varepsilon & \text{on } B_{\eta}^{c} \end{cases}, \qquad h = \begin{cases} \Phi((\alpha + \beta)\Phi^{-1}(\rho)) & \text{on } \alpha A_{\eta} + \beta B_{\eta} \\ \delta_{\varepsilon,\rho} & \text{on } (\alpha A_{\eta} + \beta B_{\eta})_{\eta}^{c} \end{cases}$$

with intermediate values elsewhere. Since $2\varepsilon < \rho$ we immediately get $\delta_{\varepsilon,\rho} \leq \Phi((\alpha+\beta)\Phi^{-1}(\rho))$, which gives $h \geq \delta_{\varepsilon,\rho}$. Since A, B are compact, for big enough a the functions f, g are equal ε outside B(0, a). Thus, for big a we have $(f, g, h) \in \mathcal{N}_{a,\varepsilon,\rho}$.

We shall verify that

$$\Phi^{-1}(h(\alpha x + \beta y)) \ge \alpha \Phi^{-1}(f(x)) + \beta \Phi^{-1}(g(y)).$$

If $x \in A_{\eta}$ and $y \in B_{\eta}$ we get $\alpha x + \beta y \in \alpha A_{\eta} + \beta B_{\eta}$ and thus

$$\alpha \Phi^{-1}(f(x)) + \beta \Phi^{-1}(g(y)) \le \alpha \Phi^{-1}(\rho) + \beta \Phi^{-1}(\rho) = \Phi^{-1}\Phi((\alpha + \beta)\Phi^{-1}(\rho)) = h(\alpha x + \beta y).$$

since always $f(x), g(y) \leq \rho$. If $x \notin A_{\eta}$ or $y \notin A_{\eta}$ then

$$\begin{aligned} \alpha \Phi^{-1}(f(x)) + \beta \Phi^{-1}(g(y)) &\leq \max \left\{ \Phi(\alpha \Phi^{-1}(\varepsilon) + \beta \Phi^{-1}(\rho)), \Phi(\alpha \Phi^{-1}(\rho) + \beta \Phi^{-1}(\varepsilon)) \right\} \\ &\leq \delta_{\varepsilon,\rho} \leq h(\alpha x + \beta y), \end{aligned}$$

since always $h \geq \delta_{\varepsilon,\rho}$. Using Theorem 24 we get

$$\Phi^{-1}\left(\int h \mathrm{d}\gamma_n\right) \ge \alpha \Phi^{-1}\left(\int f \mathrm{d}\gamma_n\right) + \beta \Phi^{-1}\left(\int g \mathrm{d}\gamma_n\right) \ge \alpha \Phi^{-1}\left(\rho\gamma_n(A)\right) + \beta \Phi^{-1}\left(\rho\gamma_n(B)\right).$$

Taking $\varepsilon \to 0^+$ and $\eta \to 0^+$ we get $\delta_{\varepsilon,\rho} \to 0$ and thus

$$\Phi^{-1}\left(\int h \mathrm{d}\gamma_n\right) \to \Phi^{-1}\left(\Phi((\alpha+\beta)\Phi^{-1}(\rho))\gamma_n(\alpha A+\beta B)\right).$$

Therefore

$$\Phi^{-1}\left(\Phi((\alpha+\beta)\Phi^{-1}(\rho))\gamma_n(\alpha A+\beta B)\right) \ge \alpha\Phi^{-1}\left(\rho\gamma_n(A)\right)+\beta\Phi^{-1}\left(\rho\gamma_n(B)\right).$$

Now we take $\rho \to 1^-$ and observe that then $\Phi((\alpha + \beta)\Phi^{-1}(\rho)) \to 1$ and thus we arrive at the desired inequality

$$\Phi^{-1}(\gamma_n(\alpha A + \beta B)) \ge \alpha \Phi^{-1}(\gamma_n(A)) + \beta \Phi^{-1}(\gamma_n(B)).$$

We shall now prove Theorem 24 for nice triples of functions (f, g, h).

Proof of Theorem 24 for $\mathcal{N}_{a,\varepsilon,\rho}$.

Step 1. We derive a PDE for Q_t . It is not surprising that we shall get the heat equation. Indeed, integration by parts gives

$$\begin{aligned} \frac{\partial}{\partial t}Q_t f(x) &= \frac{\partial}{\partial t} \int f(x + \sqrt{t}z) \mathrm{d}\gamma_n(z) = \frac{1}{2} \int \frac{1}{\sqrt{t}} \nabla f(x + \sqrt{t}z) \cdot z \mathrm{d}\gamma_n(z) \\ &= -\frac{1}{2} \int \frac{1}{\sqrt{t}} \nabla f(x + \sqrt{t}z) \cdot \nabla \varphi_n(z) \mathrm{d}z = \frac{1}{2} \int \Delta f(x + \sqrt{t}z) \varphi_n(z) \mathrm{d}z \\ &= \frac{1}{2} \Delta \int f(x + \sqrt{t}z) \mathrm{d}\gamma_n(z) = \frac{1}{2} \Delta Q_t f(x). \end{aligned}$$

Step 2. Suppose u = u(t, x) satisfies $\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u$. We would like to derive an equation for $U = \Phi^{-1}(u)$. We have $u = \Phi(U)$. Thus,

$$\frac{\partial u}{\partial t} = \Phi'(U) \frac{\partial U}{\partial t} = \varphi(U) \frac{\partial U}{\partial t}$$

and

$$\nabla u = \varphi(U) \nabla U, \qquad \Delta u = \varphi(U) \Delta U + \varphi'(U) |\nabla U|^2 = \varphi(U) \Delta U - U\varphi(U) |\nabla U|^2.$$

We get

$$\varphi(U)\nabla U = \frac{\partial u}{\partial t} = \frac{1}{2}\Delta u = \frac{1}{2}\left(\varphi(U)\Delta U - U\varphi(U)|\nabla U|^2\right).$$

Cancelling $\varphi(U) > 0$ we get

$$\frac{\partial U}{\partial t} = \frac{1}{2} \left(\Delta U - U |\nabla U|^2 \right).$$

Step 3. Define

$$C(t, x, y) = \Phi^{-1}(Q_t h(\alpha x + \beta y)) - \alpha \Phi^{-1}(Q_t f(x)) - \beta \Phi^{-1}(Q_t g(y)).$$

Since Q_0 is the identity operator, our assumption reads $C(0, x, y) \ge 0$ for all x, y and the assertion is $C(t, x, y) \ge 0$ for all x, y and $t \ge 0$. The idea is now to derive certain evolutionary

equation satisfied by C and prove an appropriate maximum principle for this equation. To simplify our notation we will be using

$$\begin{split} F &= F(t,x) = \Phi^{-1}(Q_t f(x)) \\ G &= G(t,y) = \Phi^{-1}(Q_t g(y)) \\ H &= H(t,\alpha x + \beta y) = \Phi^{-1}(Q_t h(\alpha x + \beta y)). \end{split}$$

Let us remember that in the upcoming computations F and all its derivatives will always be evaluated at (t, x), G and all its derivatives at (t, y) and H together with all its derivatives at $(t, \alpha x + \beta y)$. Clearly F does not depend on y and G does not depend on x. We have

$$C = H - \alpha F - \beta G.$$

We have

$$\nabla_x C = \alpha (\nabla H - \nabla F)$$
$$\nabla_y C = \beta (\nabla H - \nabla G).$$

and

$$\Delta_x C = \alpha^2 \Delta H - \alpha \Delta F$$
$$\Delta_y C = \beta^2 \Delta H - \beta \Delta G.$$

Moreover,

$$\sum_{1 \le i \le n} \frac{\partial^2 C}{\partial x_i \partial y_i} = \alpha \beta \Delta H.$$

Let us define the operator

$$L = \frac{1}{2} \left(\Delta_x + \Delta_y + \frac{1 - \alpha^2 - \beta^2}{\alpha \beta} \sum_{1 \le i \le n} \frac{\partial^2}{\partial x_i \partial y_i} \right).$$

Clearly,

$$LC = \frac{1}{2}(\Delta H - \alpha \Delta F - \beta G).$$

From Step 2 we get

$$\begin{split} \frac{\partial F}{\partial t} &= \frac{1}{2} \Delta F - \frac{1}{2} F |\nabla F|^2 \\ \frac{\partial G}{\partial t} &= \frac{1}{2} \Delta G - \frac{1}{2} G |\nabla G|^2 \\ \frac{\partial H}{\partial t} &= \frac{1}{2} \Delta H - \frac{1}{2} H |\nabla H|^2. \end{split}$$

24

Thus,

$$\begin{split} LC &= \frac{\partial H}{\partial t} + \frac{1}{2} H |\nabla H|^2 \\ &- \alpha \frac{\partial F}{\partial t} - \frac{1}{2} \alpha F |\nabla F|^2 \\ &- \beta \frac{\partial G}{\partial t} - \frac{1}{2} \beta G |\nabla G|^2 \\ &= \frac{\partial C}{\partial t} + \Psi, \end{split}$$

where

$$\Psi = \frac{1}{2} \left(H |\nabla H|^2 - \alpha F |\nabla F|^2 - \beta G |\nabla G|^2 \right).$$

Now

$$|\nabla F|^{2} = |\nabla H|^{2} + (\nabla F - \nabla H) \cdot (\nabla F + \nabla H) = |\nabla H|^{2} - \frac{1}{\alpha} \nabla_{x} C \cdot (\nabla F + \nabla H)$$
$$|\nabla G|^{2} = |\nabla H|^{2} + (\nabla G - \nabla H) \cdot (\nabla G + \nabla H) = |\nabla H|^{2} - \frac{1}{\beta} \nabla_{y} C \cdot (\nabla G + \nabla H)$$

Thus we can rewrite Ψ as

$$\begin{split} \Psi &= \frac{1}{2} |\nabla H|^2 (H - \alpha F - \beta G) + \frac{1}{2} \left(\nabla_x C \cdot (\nabla F + \nabla H) F + \nabla_y C \cdot (\nabla G + \nabla H) G \right) \\ &= \frac{1}{2} |\nabla H|^2 C + \frac{1}{2} \left(\nabla_x C \cdot (\nabla F + \nabla H) F + \nabla_y C \cdot (\nabla G + \nabla H) G \right) \\ &= \frac{1}{2} |\nabla H|^2 C + \nabla_{x,y} C \cdot \Theta, \end{split}$$

where Θ is a vector field in \mathbb{R}^{2n} given by

$$\Theta = \frac{1}{2}((\nabla F + \nabla H)F, (\nabla G + \nabla H)G).$$

We arrive at

(10)
$$LC = \frac{\partial C}{\partial t} + \frac{1}{2} |\nabla H|^2 C + \nabla_{x,y} C \cdot \Theta.$$

Step 4. We now show that C attains its infimum on the sets of the form $[0,T] \times \mathbb{R}^n \times \mathbb{R}^n$. Recall that we dealing with the class $\mathcal{N}_{a,\varepsilon,\rho}$. Fix T > 0. Take r such that $\gamma_n(B(0,r)) = 1 - \varepsilon$ and define $R = a + r\sqrt{T}$. If |x| > R and $t \leq T$ then for $|z| \leq r$ we have $x + \sqrt{t}z \notin B(0,a)$ and thus $f(x + \sqrt{t}z) \leq \varepsilon$, so for $t \in [0,T]$ we get

$$(Q_t f)(x) = \int_{|z| \le r} f(x + \sqrt{t}z) d\gamma_n(z) + \int_{|z| > r} f(x + \sqrt{t}z) d\gamma_n(z)$$

$$\le \varepsilon (1 - \varepsilon) + \varepsilon \rho < 2\varepsilon.$$

By the same argument $(Q_t g)(y) < 2\varepsilon$. Since $h \ge \delta_{\varepsilon,\rho}$ then also $Q_t h \ge \delta_{\varepsilon,\rho}$. Thus if |x| > R or |y| > R then either $(Q_t f)(x) < 2\varepsilon$ or $(Q_t g)(y) < 2\varepsilon$ and thus (since $Q_t f(x), Q_t g(y) \le \rho$ as $f, g \le \rho$)

$$\begin{aligned} \alpha Q_t f(x) + \beta Q_t g(y) &\leq \max \left\{ \Phi(\alpha \Phi^{-1}(2\varepsilon) + \beta \Phi^{-1}(\rho)), \Phi(\alpha \Phi^{-1}(\rho) + \beta \Phi^{-1}(2\varepsilon)) \right\} \\ &= \delta_{\varepsilon,\rho} \leq Q_t h(\alpha x + \beta y). \end{aligned}$$

Thus C is non-negative on $[0,T] \times (B(0,R) \times B(0,R))^c$.

Step 5. Suppose C(t, x, y) is negative at some point (T, x, y). From Step 4 we know that

$$0 > \inf_{[0,T] \times \mathbb{R}^n \times \mathbb{R}^n} C(t, x, y) = \inf_{[0,T] \times B \times B} C(t, x, y),$$

where B = B(0, R). Since C is continuous (in fact C is C^{∞} smooth with all partial derivatives of all orders uniformly bounded) we get that on the set $[0, T] \times B \times B$ the function C attains its negative infimum -M. Define $C_{\theta}(t, x, y) = C(t, x, y) + t\theta$ where $\theta = M/2T$. We have

$$\inf_{[0,T]\times\mathbb{R}^n\times\mathbb{R}^n} C_{\theta}(t,x,y) \leq -M + T \cdot M/2T = -M/2$$

and since $C_{\theta} \geq C$ we see that C_{θ} on the set $[0,T] \times \mathbb{R}^n \times \mathbb{R}^n$ attains its infimum in certain point $(t_0, x_0, y_0) \in [0,T] \times B \times B$ and in that point $C(t_0, x_0, y_0) \leq C_{\theta}(t_0, x_0, y_0) < 0$.

Due to our assumption $C_{\theta}(0, x, y) \ge 0$ and so the infimum is not attained on $\{0\} \times B \times B$. Also the infimum in not attained on $[0, T] \times (\partial B \times B \cup B \times \partial B)$ since on these points the function C_{θ} is non-negative. So, the infimum is attained on $(0, T) \times int(B) \times int(B)$ or on $\{T\} \times int(B) \times int(B)$.

Let us analyse the first case. At the minimal values we have

$$\nabla_{x,y}C_{\theta} = 0, \qquad \frac{\partial C_{\theta}}{\partial t} = 0, \qquad \text{Hess}_{x,y}C_{\theta} \ge 0, \qquad C_{\theta} < 0,$$

which gives

$$\nabla_{x,y}C = 0, \qquad \frac{\partial C}{\partial t} = -\theta < 0, \qquad \text{Hess}_{x,y}C \ge 0, \qquad C < 0,$$

We shall soon verify

claim:
$$\operatorname{Hess}_{x,y} C \ge 0 \implies LC \ge 0.$$

Using (10) we get

$$0 \le LC = \frac{\partial C}{\partial t} + \frac{1}{2} |\nabla H|^2 C + \nabla_{x,y} C \cdot \Theta \le -\theta,$$

which is a contradiction.

If the infimium of C_{θ} is attained on $\{T\} \times \operatorname{int}(B) \times \operatorname{int}(B)$ then the same equations for the critical point are satisfied, except for equations $\frac{\partial C_{\theta}}{\partial t} = 0$ which now has to be replaced by the inequality $\frac{\partial C_{\theta}}{\partial t} \leq 0$, leading to the same contradiction.

Step 6. We shall verify the claim. Let A be a $2n \times 2n$ matrix with n diagonal 2×2 blocks

$$A_0 = \left[\begin{array}{cc} 1 & \frac{1-\alpha^2-\beta^2}{2\alpha\beta} \\ \frac{1-\alpha^2-\beta^2}{2\alpha\beta} & 1 \end{array} \right].$$

It is straightforward to observe that

$$2LC = [1, \dots, 1] \cdot (A * \operatorname{Hess}_{x,y} C) \cdot \begin{bmatrix} 1\\ \vdots\\ 1 \end{bmatrix},$$

where * denotes the Hadamard product of two matrices, namely $(A * B) = (a_{ij}b_{ij})_{ij}$. It is therefore enough to verify that $A * \operatorname{Hess}_{x,y} C$ is positive semi-definite. Since $\operatorname{Hess}_{x,y} C$ itself is positive semi-definite and the Hadamard product of two positive semi-definite matrices (see semi-definite.

lemma below) is positive semi-definite, it is enough to verify that A is positive semi-definite, which amounts to proving that A_0 is positive semi-definite. The matrix

$$A_0(d) = \left[\begin{array}{cc} 1 & d \\ d & 1 \end{array} \right]$$

is positive semi-definite iff $|d| \leq 1$, which leads to the condition $|1 - \alpha^2 - \beta^2| \leq 2\alpha\beta$. This is

$$-2\alpha\beta \le 1 - \alpha^2 - \beta^2 \le 2\alpha\beta,$$

which is $(\alpha + \beta)^2 \ge 1$ and $(\alpha - \beta)^2 \le 1$. These are the assumptions of Theorem 24. \Box Lemma 8. Hadamard product of two symmetric positive semi-definite matrices is positive

Proof. Let $A = (a_{ij})_{i,j=1}^n$ and $B = (b_{ij})_{i,j=1}^n$. We have $A * B = (a_{ij}b_{ij})_{i,j=1}^n$. Since B is symmetric and positive semi-definite, where is an orthogonal matrix U such that $B = UDU^T$ where $D = \text{diag}(\lambda_1, \ldots, \lambda_n)$ is diagonal. All the entries of D (eigenvalues of B) are nonnegative as B was positive semi-definite. Let $\sqrt{D}\text{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$. Define $C = U\sqrt{D}U^T$. We have $B = C^2$. Thus if $v = (v_1, \ldots, v_n)$ then

$$\langle (A*B)v,v\rangle = \sum_{i,j} v_i a_{ij} b_{ij} v_j = \sum_{i,j} v_i a_{ij} b_{ji} v_j = \sum_{i,j,k} v_i a_{ij} c_{jk} c_{ki} v_j = \sum_{i,j,k} c_{ki} v_i a_{ij} v_j c_{jk}.$$

The last expression in equal to tr(CVAVC), where $V = diag(v_1, \ldots, v_n)$. The operation $A \rightarrow SAS$ preserves positive semi-definiteness as for any vector u we have (setting w = Su)

$$\langle SASu, u \rangle = \langle ASu, Su \rangle = \langle Aw, w \rangle \ge 0.$$

It suffices to use this fact twice for S = V and S = C.

We now formulate some corollaries of the Ehrhard inequality.

Corollary 25. Let K be convex in \mathbb{R}^n . Then for all $t \ge 1$ we have

$$\Phi^{-1}(\gamma_n(tK)) \ge t\Phi^{-1}(\gamma_n(K))$$

Proof. Using Ehrhard inequality with A = B = K and $\alpha = \beta = t/2$ yields the desired inequality. Note that $\frac{t}{2}K + \frac{t}{2}K = tK$ due to the convexity of K.

Corollary 26. Let A be Borel and let K be convex. Suppose $\alpha, \beta > 0, \alpha + \beta \ge 1$ and $\alpha - \beta \le 1$. Then we have

$$\Phi^{-1}(\gamma_n(\alpha A + \beta K)) \ge \alpha \Phi^{-1}(\gamma_n(A)) + \beta \Phi^{-1}(\gamma_n(K)).$$

Proof. If additionally $\beta - \alpha \leq 1$ then the assumptions of the Ehrhard inequality are satisfied, so the desired inequality follows. Suppose that $\frac{\beta}{\alpha+1} > 1$. Let us use Ehrhard inequality with α and $\beta = 1 + \alpha$ with the sets A and $\frac{\beta}{\alpha+1}K$. We get

$$\Phi^{-1}(\gamma_n(\alpha A + \beta K)) = \Phi^{-1}\left(\gamma_n(\alpha A + (\alpha + 1) \cdot \frac{\beta}{\alpha + 1}K)\right)$$

$$\geq \alpha \Phi^{-1}(\gamma_n(A)) + (\alpha + 1)\Phi^{-1}\left(\gamma_n\left(\frac{\beta}{\alpha + 1}K\right)\right).$$

Now it suffices to use Corollary 25 with $t = \frac{\beta}{\alpha+1} > 1$ to get

$$(\alpha + 1)\Phi^{-1}\left(\gamma_n\left(\frac{\beta}{\alpha + 1}K\right)\right) \ge \beta\Phi^{-1}\left(\gamma_n\left(K\right)\right)$$

Our last application is the Gaussian isoperimetric inequality. Before we prove it let us establish the following simple lemma.

Lemma 9. We have

$$\sup_{r>0} \frac{1}{r} \Phi^{-1}(\gamma(rB_2^n)) = \lim_{r \to \infty} \frac{1}{r} \Phi^{-1}(\gamma(rB_2^n)) = 1.$$

Proof. We have , by de l'Hospital

$$1 - \Phi(r) = \frac{1}{\sqrt{2\pi}} \int_{r}^{\infty} e^{-s^{2}/2} \mathrm{d}s \sim_{r \to \infty} \frac{1}{\sqrt{2\pi}} \frac{1}{r} e^{-r^{2}/2}.$$

Thus $\ln(1-\Phi(r)) \sim_{r\to\infty} -r^2/2$ and therefore, taking $x = \Phi(r) \to 1$, we get $\Phi^{-1}(x) \sim_{x\to 1} \sqrt{-2\ln(1-x)}$. As a consequence $\Phi^{-1}(\gamma(rB_2^n)) \sim_{r\to\infty} \sqrt{-2\ln(1-\gamma_n(rB_2^n))} \sim_{r\to\infty} r$ since

$$1 - \gamma_n(rB_2^n) = \frac{n|B_2^n|}{\sqrt{2\pi^n}} \int_r^\infty s^{n-1} e^{-s^2/2} \mathrm{d}s \sim_{r \to \infty} \frac{n|B_2^n|}{\sqrt{2\pi^n}} r^{n-2} e^{-r^2/2},$$

again by de l'Hospital. This shows the second inequality.

Now it suffices to observe that

$$\Phi^{-1}(\gamma_n(rB_2^n)) \le \Phi^{-1}(\{x_1 \le r\}) = \Phi^{-1}(\Phi(r)) = r.$$

Proof of Gaussian isoperimetry. By Corollary 26 we get

$$\Phi^{-1}(\gamma_n(A_{\varepsilon})) = \Phi^{-1}\left(\gamma_n\left(A + \frac{\varepsilon}{r} \cdot rB_2^n\right)\right)$$

$$\geq \Phi^{-1}(\gamma_n(A)) + \frac{\varepsilon}{r}\Phi^{-1}(\gamma_n(rB_2^n)) \xrightarrow[r \to \infty]{} \Phi^{-1}(\gamma_n(A)) + \varepsilon.$$

1.12. **Brascamp-Lieb inequality.** In this section we will be using the following notation. Let $S^+(\mathbb{R}^n)$ be $n \times n$ positive definite matrices. For a $m \times n$ matrix B we denote by B^* its $n \times n$ transposition. The space of all linear maps $\mathbb{R}^n \to \mathbb{R}^{n_i}$ will be denoted by $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^{n_i})$. If $A \in S^+(\mathbb{R}^k)$ then $G_A(x) = \exp(-\langle Ax, x \rangle)$ be the corresponding Gaussian function. Note that

$$\int_{\mathbb{R}^k} G_A(x) \mathrm{d}x = \pi^{k/2} \det(A)^{-1/2}$$

Let $L_1^+(\mathbb{R}^n)$ be the space of non-negative functions from $L_1(\mathbb{R}^n)$. For a function $m : \mathbb{R}^n \to \mathbb{R}_+$ we put

$$\int^{\downarrow} m = \sup \left\{ \int \tilde{m} : \ \tilde{m} \le m, \ m \text{ is measurable} \right\}.$$

Theorem 27. Let $m \ge n$ be positive integers and let $c_1, \ldots, c_m > 0$ be real numbers. Let $n_1, \ldots, n_m \le n$ be positive integers such that $n = \sum_{i=1}^m c_i n_i$. Suppose $B_i \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^{n_i})$ be surjective. Assume $\bigcap_{i=1}^m \ker(B_i) = \{0\}$. For $f_i \in L_1^+(\mathbb{R}^{n_i})$, $i = 1, \ldots, m$ set us define

$$J(f_1,\ldots,f_m) = \int_{\mathbb{R}^n} \prod_{i=1}^m f_i^{c_i}(B_i x) \mathrm{d}x$$

and

28

$$I(f_1, \dots, f_m) = \int_{\mathbb{R}^n}^{\downarrow} \sup \left\{ \prod_{i=1}^m f_i^{c_i}(y_i) : \quad x = \sum_{i=1}^m c_i B_i^* y_i, \quad y_i \in \mathbb{R}^{n_i} \right\} \mathrm{d}x$$

Let E, F be best constants in the inequalities

$$I(f_1,\ldots,f_m) \ge E \cdot \prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} f_i \right)^{c_i}, \qquad J(f_1,\ldots,f_m) \le F \cdot \prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} f_i \right)^{c_i}$$

Let

$$E_{g} = \int \left\{ \frac{I(G_{A_{1}}, \dots, G_{A_{m}})}{\prod_{i=1}^{m} \left(\int_{\mathbb{R}^{n_{i}}} G_{A_{i}} \right)^{c_{i}}} : A_{i} \in S^{+}(\mathbb{R}^{n_{i}}), \ i = 1, \dots, m \right\},$$

$$F_{g} = \sup \left\{ \frac{J(G_{A_{1}}, \dots, G_{A_{m}})}{\prod_{i=1}^{m} \left(\int_{\mathbb{R}^{n_{i}}} G_{A_{i}} \right)^{c_{i}}} : A_{i} \in S^{+}(\mathbb{R}^{n_{i}}), \ i = 1, \dots, m \right\},$$

be the corresponding best constants for Gaussian functions. Let D be the best constant in the inequality

(11)
$$\det\left(\sum_{i=1}^{m} c_i B_i^* A_i B_i\right) \ge D \prod_{i=1}^{m} \det(A_i)^{c_i}.$$

Then

$$E = E_g = \sqrt{D}$$
, and $F = F_g = \frac{1}{\sqrt{D}}$.

Remark 1. The condition $\bigcap_{i=1}^{m} \ker(B_i) = \{0\}$ ensures that $\sum_{i=1}^{m} c_i B_i^* A_i B_i$ is non-singular. Otherwise the theorem still holds true with D = 0.

Remark 2. The condition $n = \sum_{i=1}^{m} c_i n_i$ is the condition needed for the homogeneity of (11) under $A_i \to \lambda A_i$.

Remark 3. By using approximations similar to those described in the transportation proof of Prékopa-Leindler we can assume that the functions f_1, \ldots, f_m are of the form $f_i = \tilde{f}_i \mathbf{1}_{\Omega_i}$, where Ω_i is some open Euclidean ball in \mathbb{R}^{n_i} , the function \tilde{f}_i is Lipschitz and satisfies $0 < c_i \leq \tilde{f}_i \leq C_i$ for some positive finite constants c_i, C_i . We shall denote this class of functions f_i by $C_L(\mathbb{R}^{n_i})$. The proof of Theorem 27 uses the following regularity theorem.

Theorem 28. Suppose $f, h \in C_L(\mathbb{R}^n)$ be probability densities with open domains Ω_f, Ω_h and let μ_f, μ_g be the corresponding probability measures. Then there is a $C^2(\Omega_h)$ convex function ϕ such that $\mu_h = T\mu_f$, where $T = \nabla \phi$. Moreover, the following transport equation is satisfied,

$$\det(DT(x))f(Tx) = h(x).$$

Lemma 10. We have $F_g = \frac{1}{\sqrt{D}}$.

Proof. We have

$$J(G_{A_1,\dots},G_{A_m}) = \int_{\mathbb{R}^n} \prod_{i=1}^m \exp(-c_i \langle A_i B_i x, B_i x \rangle) dx = \int_{\mathbb{R}^n} \exp\left(-\sum_{i=1}^m c_i \langle A_i B_i x, B_i x \rangle\right) dx$$
$$= \int_{\mathbb{R}^n} \exp\left(-\left\langle\sum_{i=1}^m c_i B_i^* A_i B_i x, x\right\rangle\right) dx = \pi^{n/2} \det\left(\sum_{i=1}^m c_i B_i^* A_i B_i\right)^{-1/2}.$$

$$\prod_{i=1}^{m} \left(\int_{\mathbb{R}^{n_i}} G_{A_i} \right)^{c_i} = \prod_{i=1}^{m} \pi^{c_i n_i/2} \det(A_i)^{-c_i/2}$$
$$= \pi^{\frac{1}{2} \sum_{i=1}^{m} c_i n_i} \prod_{i=1}^{m} \det(A_i)^{-c_i/2} = \pi^{n/2} \prod_{i=1}^{m} \det(A_i)^{-c_i/2}.$$

Thus

$$F_g = \sup_{A_1,\dots,A_m} \left(\frac{\det\left(\sum_{i=1}^m c_i B_i^* A_i B_i\right)}{\prod_{i=1}^m \det(A_i)^{c_i}} \right)^{-1/2} = \left(\inf_{A_1,\dots,A_m} \frac{\det\left(\sum_{i=1}^m c_i B_i^* A_i B_i\right)}{\prod_{i=1}^m \det(A_i)^{c_i}} \right)^{-1/2} = D^{-1/2}.$$

Lemma 11. We have $E_g F_g = 1$.

Proof. Let

$$Q = \sum_{i=1}^{m} c_i B_i^* A_i B_i$$
, and $Q(y) = \langle Qy, y \rangle$,

where we slightly abused notation. This matrix is symmetric positive definite since

$$\langle Qy, y \rangle = \sum_{i=1}^{m} c_i \langle A_i B_i v, B_i v \rangle \ge 0$$

and its is equal zero only if $B_i v = 0$ for i = 1, ..., m, which means that $v \in \bigcap_{i=1}^m \ker(B_i) = \{0\}$ by our assumptions. In particular $\det(Q) > 0$. We saw in the proof of Lemma 10 that

(12)
$$\frac{J(G_{A_1},\ldots,G_{A_m})}{\prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} G_{A_i}\right)^{c_i}} = \left(\frac{\prod_{i=1}^m \det(A_i)^{c_i}}{\det(Q)}\right)^{1/2}.$$

We define the dual of this quadratic form as

$$Q_*(x) = \sum \{ |\langle x, y \rangle| : Q(y) \le 1 \}.$$

Claim 1. $Q_*(x) = \langle Q^{-1}x, x \rangle.$

Proof of Claim 1. We observe that

$$|\langle x,y\rangle|^2 \le \langle Q^{-1}x,x\rangle\langle Qy,y\rangle$$

with equality for $y = Q^{-1}x$. Indeed, we have

$$|\langle x, y \rangle|^{2} = |\langle Q^{-1/2}x, Q^{1/2}y \rangle|^{2} \le |Q^{-1/2}x|^{2} \cdot |Q^{1/2}y|^{2} = \langle Q^{-1}x, x \rangle \langle Qy, y \rangle.$$

Here we have used the fact that

$$|Q^{1/2}y|^2 = \langle Q^{1/2}y, Q^{1/2}y \rangle = \langle Q^{1/2}Q^{1/2}y, y \rangle = \langle Qy, y \rangle.$$

Now we define

$$R(x) = \int \left\{ \sum_{i=1}^{m} c_i \left\langle A_i^{-1} y_i, y_i \right\rangle : \ x = \sum_{i=1}^{m} c_i B_i^* y_i, \ y_i \in \mathbb{R}^{n_i}, \ i = 1, \dots, m \right\}.$$

Note that R is highly relevant to the computation of $I(f_1, \ldots, f_m)$, namely

$$I(G_{A_1^{-1}}, \dots, G_{A_m^{-1}}) = \int_{\mathbb{R}^n} \exp(-R(x)) \mathrm{d}x.$$

Claim 2. We have $R = Q_*$. In particular, using Claim 1 we get $R(x) = \langle Q^{-1}x, x \rangle$. *Proof of Claim 2.* Suppose $x = \sum_{i=1}^m c_i B_i^* y_i$, for some $y_i \in \mathbb{R}^{n_i}$. Then we have

$$|\langle x, y \rangle|^{2} = \left| \left\langle \sum_{i=1}^{m} c_{i} B_{i}^{*} y_{i}, y \right\rangle \right|^{2} = \left| \sum_{i=1}^{m} \left\langle \sqrt{c_{i}} y_{i}, \sqrt{c_{i}} B_{i} y \right\rangle \right|^{2} = \left| \sum_{i=1}^{m} \left\langle \sqrt{c_{i}} A_{i}^{-1/2} y_{i}, \sqrt{c_{i}} A_{i}^{1/2} B_{i} y \right\rangle \right|^{2}$$

$$= \left| \sum_{i=1}^{m} |\sqrt{c_{i}} A_{i}^{-1/2} y_{i}| \cdot |\sqrt{c_{i}} A_{i}^{1/2} B_{i} y| \right|^{2} \leq \left(\sum_{i=1}^{m} |\sqrt{c_{i}} A_{i}^{-1/2} y_{i}|^{2} \right) \left(\sum_{i=1}^{m} |\sqrt{c_{i}} A_{i}^{1/2} B_{i} y|^{2} \right)$$

$$= \left(\sum_{i=1}^{m} c_{i} \left\langle y_{i}, A_{i}^{-1} y_{i} \right\rangle \right) \left\langle \sum_{i=1}^{m} c_{i} B_{i}^{*} A_{i} B_{i} y, y \right\rangle = \left(\sum_{i=1}^{m} c_{i} \left\langle A_{i}^{-1} y_{i}, y_{i} \right\rangle \right) Q(y)$$

Taking the infimum with respect to y_i gives

$$|\langle x, y \rangle|^2 \le R(x)Q(y).$$

In particular $R \ge Q_*$. To see that we actually have equality it suffices to show that for every fixed x there is y with $|\langle x, y \rangle|^2 = R(x)Q(y)$. Take

$$y = Q^{-1}x$$
, and $x_i = A_i B_i y$.

We have

$$R(x) \le \sum_{i=1}^{m} c_i \left\langle x_i, A_i^{-1} x_i \right\rangle = \sum_{i=1}^{m} c_i \left\langle A_i B_i y, B_i y \right\rangle = \sum_{i=1}^{m} c_i \left\langle B_i^* A_i B_i y, y \right\rangle = Q(y)$$

and thus

$$|\langle x, y \rangle|^2 = |\langle Qy, y \rangle|^2 = Q(y)^2 \ge R(x)Q(y) \ge |\langle x, y \rangle|^2,$$
so we must have $R(x)Q(y) = |\langle x, y \rangle|^2$.

Using Claim 2 we get

$$I(G_{A_1^{-1}}, \dots, G_{A_m^{-1}}) = \int_{\mathbb{R}^n} \exp(-R(x)) dx = \int_{\mathbb{R}^n} \exp(-\langle Q^{-1}x, x \rangle) dx$$
$$= \pi^{n/2} \det(Q^{-1})^{-1/2} = \pi^{n/2} \det(Q)^{1/2}.$$

Therefore,

$$\frac{I(G_{A_1^{-1}},\ldots,G_{A_m^{-1}})}{\prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} G_{A_i^{-1}}\right)^{c_i}} = \frac{\pi^{n/2} \det(Q)^{1/2}}{\prod_{i=1}^m \pi^{c_i n_i/2} \det(A_i^{-1})^{c_i/2}} = \det(Q)^{1/2} \prod_{i=1}^m \det(A_i)^{-c_i/2}.$$

Combining this with (12) gives

$$\frac{J(G_{A_1},\ldots,G_{A_m})}{\prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} G_{A_i}\right)^{c_i}} = \left(\frac{I(G_{A_1^{-1}},\ldots,G_{A_m^{-1}})}{\prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} G_{A_i^{-1}}\right)^{c_i}}\right)^{-1}$$

Taking the supremum of both sided gives $F_g = 1/E_g$.

We shall also need the following lemma.

Lemma 12. Suppose $f_i, h_i \in C_L(\mathbb{R}^{n_i}), i = 1, ..., m$ satisfy $\int_{\mathbb{R}^{n_i}} f_i = \int_{\mathbb{R}^{n_i}} h_i = 1$. Then $I(f_1, ..., f_m) \geq DJ(h_1, ..., h_m).$

Proof. We can assume D > 0, since otherwise there is nothing there to prove. Recall that h_i are restriction of some positive Lipschitz function to some open domains Ω_{h_i} . Take the Brenier maps T_i transporting the probability measure with density h_i onto the probability measure with density f_i . By the Brenier-Cafarelli theorem we have $T_i = \nabla \phi_i$, where ϕ_i is a convex $C^2(\Omega_h)$ function. Thus, DT_i is positive semi-definite. Moreover, we have the following transport equation

$$\det(DT_i(x))f_i(T_ix) = h_i(x).$$

Since $h_i(x) > 0$ on Ω_{h_i} , we get that $\det(DT_i) > 0$ and thus DT_i is positive definite. Define $S = \bigcap_{i=1}^m B_i^{-1}(\Omega_{h_i}) \subseteq \mathbb{R}^n$ and consider $\Theta : S \to \mathbb{R}^n$ given by

$$\Theta(y) = \sum_{i=1}^{m} c_i B_i^* T_i(B_i y).$$

We have

$$D\Theta(y) = \sum_{i=1}^{m} c_i B_i^* DT_i(B_i y) B_i.$$

Thus, $D\Theta(y)$ is positive definite (we argue similarly to the proof of Lemma 11). alternatively we could write that

$$\det(D\Theta(y)) = \det\left(\sum_{i=1}^{m} c_i B_i^* DT_i(B_i y) B_i\right) \ge D \prod_{i=1}^{m} \det(DT_i(B_i y))^{c_i} > 0.$$

Thus, $\langle D\Theta(y)v, v \rangle$ for any $v \neq 0$ and so

$$\langle y - x, \Theta(y) - \Theta(x) \rangle = \left\langle y - x, \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} \Theta(ty + (1 - t)x) \mathrm{d}t \right\rangle$$

= $\left\langle y - x, \int_0^1 D\Theta(ty + (1 - t)x)(y - x) \mathrm{d}t \right\rangle$
= $\int_0^1 \left\langle D\Theta(ty + (1 - t)x)(y - x), y - x \right\rangle \mathrm{d}t > 0.$

Thus, Θ is an injective map.

We have

$$\begin{split} \int_{\mathbb{R}^n} \prod_{i=1}^n h_i^{c_i}(B_i y) \mathrm{d}y &= \int_S \prod_{i=1}^n h_i^{c_i}(B_i y) \mathrm{d}y = \int_S \prod_{i=1}^n \left(f_i(T_i(B_i y)) \det(DT_i(B_i y)) \right)^{c_i} \mathrm{d}y \\ &\leq \frac{1}{D} \int_S \prod_{i=1}^n \left(f_i(T_i(B_i y)) \right)^{c_i} \det\left(\sum_{i=1}^m c_i B_i^* DT_i(B_i y) B_i\right) \mathrm{d}y \\ &\leq \frac{1}{D} \int_S \left(\sup_{\Theta(y) = \sum_{i=1}^m c_i B_i^* y_i} \prod_{i=1}^m f_i(y_i)^{c_i} \right) \det(\Theta(y)) \mathrm{d}y \\ &\leq \frac{1}{D} \int_{\mathbb{R}^n} \sup_{x = \sum_{i=1}^m c_i B_i^* y_i} \prod_{i=1}^m f_i(y_i)^{c_i}. \end{split}$$

Here the last equality is a change of variables and the second inequality is

$$\prod_{i=1}^{m} (f_i(T_i(B_i y)))^{c_i} \le \sup_{\Theta(y) = \sum_{i=1}^{m} c_i B_i^* y_i} \prod_{i=1}^{m} f(y_i)^{c_i},$$

which follows from the fact that $y_i = T_i(B_i y)$ satisfies

$$\sum_{i=1}^{m} c_i B_i^* y_i = \sum_{i=1}^{m} c_i B_i^* T_i(B_i y) = \Theta(y).$$

Note that we have used the usual integral instead of the inner integral since the function

$$x \mapsto \sup_{x = \sum_{i=1}^{m} c_i B_i^* y_i} \prod_{i=1}^{m} f_i(y_i)^{c_i}$$

is measurable (due to the continuity of f_i 's on their support one can replace the sup with a supremum on the dense countable subset of the space $\{(y_1, \ldots, y_m) : x = \sum_{i=1}^m c_i B_i^* y_i\}$; see the discussion before the transportation proof of Prékopa-Leindler inequality). \Box

Proof of BL and RBL. Note that due to the invariance under scaling $f_i \rightarrow \lambda_i f_i$ we have

$$E_g \ge E = \inf\{I(g_1, \dots, g_m), g_i - \text{centered Gaussian densities}\}$$
$$\ge D \sup\{J(g_1, \dots, g_m), g_i - \text{centered Gaussian densities}\}$$
$$= DF_g = E_g,$$

where the second inequality follows from Lemma 12 and the last equality from Lemma 10 and Lemma 11. $\hfill \Box$

Example 1. Take $n_1 = \ldots = n_m = n$ and $c_i > 0$, $i = 1, \ldots, m$ such that $\sum_{i=1}^m c_i = 1$. Moreover, let $B_1 = \ldots = B_m = I$. Then the BL inequality reads

$$\int_{\mathbb{R}^n} \prod_{i=1}^m f_i(x)^{c_i} \le F \prod_{i=1}^m \left(\int_{\mathbb{R}^n} f_i \right)^{c_i},$$

which is Hölder inequality. We know that the best constant in Hölder inequality is F = 1and thus we expect that D = 1. To check it we have to prove that

$$\det(\sum_{i=1}^{m} c_i A_i) \ge \prod_{i=1}^{m} \det(A_i)^{c_i}, \qquad A_i \in S^+(\mathbb{R}^n).$$

We proceed by induction on m. Suppose we can prove it for $m \ge 2$. For the induction steps we write

$$\det\left(\sum_{i=1}^{m+1} c_i A_i\right) = \det\left(\sum_{i=1}^{m-1} c_i A_i + \left(\frac{c_m}{c_m + c_{m+1}} A_m + \frac{c_{m+1}}{c_m + c_{m+1}} A_{m+1}\right) (c_m + c_{m+1})\right)$$
$$\geq \prod_{i=1}^{m-1} \det(A_i)^{c_i} \det\left(\frac{c_m}{c_m + c_{m+1}} A_m + \frac{c_{m+1}}{c_m + c_{m+1}} A_{m+1}\right)^{c_m + c_{m+1}} \geq \prod_{i=1}^m \det(A_i)^{c_i},$$

where the last inequality follows from the case m = 2. To prove the assertion for m = 2 we assume without loss of generality that A is invertible (otherwise approximate A by invertible

matrices). Assume we know how to prove the inequality for A = I. Then

$$\det(\lambda A + (1-\lambda)B) = \det(A)\det(\lambda + (1-\lambda)A^{-1}B) \ge \det(A)\det(A^{-1}B)^{1-\lambda}$$
$$= \det(A)^{\lambda}\det(A)^{1-\lambda}\det(A^{-1}B)^{1-\lambda} = \det(A)^{\lambda}\det(B)^{1-\lambda}.$$

If A = I then by applying orthogonal transformation we can assume that B is diagonal with positive eigenvalues a_1, \ldots, a_n . Then the inequality reads

$$\prod_{i=1}^{m} (\lambda + (1-\lambda)a_i) \ge \prod_{i=1}^{m} a_i^{1-\lambda}.$$

Clearly it suffices to prove it for m = 1. then it reads $\lambda + (1 - \lambda)a \ge a^{1-\lambda}$, which is the concavity of the logarithm,

$$\log(\lambda + (1 - \lambda)a) \ge \lambda \log(1) + (1 - \lambda) \log a = (1 - \lambda) \log a = \log(a^{1 - \lambda}).$$

The RBL inequality for this choice of n_i and B_i reads as follows: Whenever $\sum_{i=1}^{m} c_i = 1$ then

$$\int_{\mathbb{R}^n}^{\downarrow} \sup\left\{\prod_{i=1}^m f_i(y_i)^{c_i}: x = \sum_{i=1}^m c_i y_i, y_i \in \mathbb{R}^n\right\} dx \ge \prod_{i=1}^m \left(\int_{\mathbb{R}^n} f_i\right)^{c_i}.$$

This is a generalization of Prékopa-Leindler inequality for the case of m functions.

Example 2. We shall prove the following theorem

Theorem 29. Suppose $p, q, r \ge 1$ with $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$. Assume $f \in L_p(\mathbb{R}^n)$ and $g \in L_q(\mathbb{R})$. Then

(13)
$$||f * g||_r \le \left(\frac{C_p C_q}{C_r}\right)^n ||f||_p ||g||_q, \qquad C_s^2 = \frac{s^{1/s}}{s'^{1/s'}}, \quad \frac{1}{s} + \frac{1}{s'} = 1.$$

To see the connection to Brascamp-Lieb inequality let us state an alternative equivalent form of the above inequality.

Theorem 30. Suppose $p, q, r \ge 1$ with $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2$. Assume $f \in L_p(\mathbb{R}^n)$, $g \in L_q(\mathbb{R}^n)$ and $h \in L_r(\mathbb{R}^n)$. Then

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-y)g(y)h(x) \mathrm{d}x \mathrm{d}y \le (C_p C_r C_r)^n \left\|f\right\|_p \left\|g\right\|_q \left\|h\right\|_r$$

To see the equivalence of Young's inequality and the above theorem we observe that

$$\sup\left\{\frac{\|f * g\|_{r}}{\|f\|_{p} \|g\|_{q}} : f \in L_{p}(\mathbb{R}^{n}), g \in L_{q}(\mathbb{R}^{n})\right\}$$
$$= \sup\left\{\frac{\int_{\mathbb{R}^{n}} \int_{\mathbb{R}^{n}} f(x - y)g(y)h(x)dxdy}{\|f\|_{p} \|g\|_{q} \|h\|_{r'}} : f \in L_{p}(\mathbb{R}^{n}), g \in L_{q}(\mathbb{R}^{n}), h \in L_{r'}(\mathbb{R}^{n})\right\}$$

Note that $\frac{1}{p} + \frac{1}{q} + \frac{1}{r'} = 1 + \frac{1}{r} + \frac{1}{r'} = 2$. Moreover $C_{r'} = 1/C_r$.

For n = 1 we also have

$$\begin{split} \sup \left\{ \frac{\int_{\mathbb{R}^{n}} \int_{\mathbb{R}^{n}} f(x-y)g(y)h(x)dxdy}{\|f\|_{p} \|g\|_{q} \|h\|_{r'}} : \ f \in L_{p}(\mathbb{R}^{n}), g \in L_{q}(\mathbb{R}^{n}), h \in L_{r'}(\mathbb{R}^{n}) \right\} \\ &= \sup \left\{ \frac{\int_{\mathbb{R}^{n}} \int_{\mathbb{R}^{n}} f^{\frac{1}{p}}(x-y)g^{\frac{1}{q}}(y)h^{\frac{1}{r'}}(x)dxdy}{\left(\int_{\mathbb{R}^{n}} f\right)^{\frac{1}{p}} \left(\int_{\mathbb{R}^{n}} g\right)^{\frac{1}{q}} \left(\int_{\mathbb{R}^{n}} h\right)^{\frac{1}{r'}} : \ f, g, h \in L_{1}^{+}(\mathbb{R}) \right\} \\ &= \sup \left\{ \frac{\int_{\mathbb{R}} \int_{\mathbb{R}} f^{\frac{1}{p}}(x,y) \cdot (1,-1)g^{\frac{1}{q}}(x,y) \cdot (0,1)h^{\frac{1}{r'}}(x,y) \cdot (1,0)dxdy}{\left(\int_{\mathbb{R}} f\right)^{\frac{1}{p}} \left(\int_{\mathbb{R}} g\right)^{\frac{1}{q}} \left(\int_{\mathbb{R}} h\right)^{\frac{1}{r'}}} : \ f, g, h \in L_{1}^{+}(\mathbb{R}) \right\}. \end{split}$$

This relates the problem to the quantity studied by Brascamp and Lieb. We can now give a proof in dimension n = 1 using BL inequality.

Proof of Young's inequality for n = 1. Suppose $p, q, r \ge 1$ with $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2$. From BL inequality the best constant F in the inequality

$$\int_{\mathbb{R}} \int_{\mathbb{R}} f^{\frac{1}{p}}(x-y) g^{\frac{1}{q}}(y) h^{\frac{1}{r}}(x) \mathrm{d}x \mathrm{d}y \le F\left(\int_{\mathbb{R}} f\right)^{\frac{1}{p}} \left(\int_{\mathbb{R}} g\right)^{\frac{1}{q}} \left(\int_{\mathbb{R}} h\right)^{\frac{1}{r}}$$

is $1/\sqrt{D}$, where D is the best constant in the inequality

$$\det\left(c_1a_1(1,-1)^T(1,-1)+c_2a_2(0,1)^T(0,1)+c_3a_3(1,0)^T(1,0)\right) \ge a_1^{c_1}a_2^{c_2}a_3^{c_3},$$

where $c_1 = 1/p$, $c_2 = 1/q$ and $c_3 = 1/r$. Note that $c_1 + c_2 + c_3 = 2$. The determinant of the matrix on the left hand side is equal to

$$\det \begin{pmatrix} c_1a_1 + c_3a_3 & -c_1a_1 \\ -c_1a_1 & c_1a_1 + c_2a_2 \end{pmatrix} = (c_1a_1 + c_3a_3)(c_1a_1 + c_2a_2) - c_1^2a_1^2$$
$$= c_1c_2a_1a_2 + c_2c_3a_2a_3 + c_3c_1a_3a_1.$$

Thus, we ask for the best constant in

$$c_1c_2a_1a_2 + c_2c_3a_2a_3 + c_3c_1a_3a_1 \ge a_1^{c_1}a_2^{c_2}a_3^{c_3}.$$

Without loss of generality we can assume that p, q, r > 1 and thus $c_1, c_2, c_3 < 1$. By AM-GM we have

$$\begin{split} c_1c_2a_1a_2 + c_2c_3a_2a_3 + c_3c_1a_3a_1 &= (1-c_3)\frac{c_1c_2a_1a_2}{1-c_3} + (1-c_1)\frac{c_2c_3a_2a_3}{1-c_1} + (1-c_2)\frac{c_3c_1a_3a_1}{1-c_2} \\ &\geq \frac{c_1c_2}{1-c_3}\frac{1-c_3}{1-c_1}\frac{c_2c_3}{1-c_1}\frac{1-c_1}{1-c_2}a_1^{2-c_2-c_3}a_2^{2-c_1-c_3}a_3^{2-c_1-c_2} \\ &= \frac{c_1c_2}{1-c_3}\frac{1-c_3}{1-c_1}\frac{c_2c_3}{1-c_1}\frac{1-c_1}{1-c_2}\frac{c_3c_1}{1-c_2}\frac{1-c_2}{1-c_2}a_1^{c_1}a_2^{c_2}a_3^{c_3}. \end{split}$$

The equality holds for a_1, a_2, a_3 such that

$$\frac{c_1c_2}{1-c_3}a_1a_2 = \frac{c_2c_3}{1-c_1}a_2a_3 = \frac{c_3c_1}{1-c_2}a_3a_1.$$

Solving this system of equations is straightforward (we leave is as an exercise). It suffices to observe that

$$\frac{c_1c_2}{1-c_3} \frac{1-c_3}{1-c_1} \frac{c_2c_3}{1-c_1} \frac{1-c_1}{1-c_2} = (p')^{\frac{1}{p'}} (q')^{\frac{1}{q'}} (r')^{\frac{1}{r'}} c_1^{2-c_2-c_3} c_2^{2-c_1-c_3} c_3^{2-c_1-c_2}$$
$$= (p')^{\frac{1}{p'}} (q')^{\frac{1}{q'}} (r')^{\frac{1}{r'}} c_1^{c_1} c_2^{c_2} c_3^{c_3} = \frac{(p')^{\frac{1}{p'}}}{p^{\frac{1}{p}}} \frac{(q')^{\frac{1}{q'}}}{q^{\frac{1}{q}}} \frac{(r')^{\frac{1}{r'}}}{r^{\frac{1}{r}}}$$
$$= (C_p C_q C_r)^{-2}.$$

We now show that Young's inequality in dimension n = 1 implies Young's inequality in any dimension. Note that $||f * g||_{\infty} \leq ||f||_r ||g||_{r'}$ and thus, with $\tilde{h}(x) = h(-x)$, we have

$$\sup_{h} \frac{\|f * g * h\|_{\infty}}{\|f\|_{p} \|g\|_{q} \|h\|_{r'}} \le \frac{\|f * g\|_{r}}{\|f\|_{p} \|g\|_{q}} = \sup_{h} \frac{\int_{\mathbb{R}^{n}} f(x - y)g(y)h(x)dxdy}{\|f\|_{p} \|g\|_{q} \|h\|_{r'}}.$$

On the other hand,

$$\begin{split} \int_{\mathbb{R}^n} f(x-y)g(y)h(x)\mathrm{d}x\mathrm{d}y &= \int_{\mathbb{R}^n} (f*g)(x)h(x)\mathrm{d}x = \int_{\mathbb{R}^n} (f*g)(x)\tilde{h}(-x)\mathrm{d}x \\ &= f*g*\tilde{h}(0) \le \left\|f*g*\tilde{h}\right\|_{\infty}. \end{split}$$

Thus

$$\sup_{h} \frac{\|f \ast g \ast h\|_{\infty}}{\|f\|_{p} \, \|g\|_{q} \, \|h\|_{r'}} \leq \frac{\|f \ast g\|_{r}}{\|f\|_{p} \, \|g\|_{q}} \leq \sup_{h} \frac{\left\|f \ast g \ast \tilde{h}\right\|_{\infty}}{\|f\|_{p} \, \|g\|_{q} \, \|h\|_{r'}} = \sup_{h} \frac{\|f \ast g \ast h\|_{\infty}}{\|f\|_{p} \, \|g\|_{q} \, \|h\|_{r'}}.$$

Therefore Young's inequality is equivalent to the following theorem.

Theorem 31. Suppose $p, q, r \ge 1$ with $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2$. Assume $f \in L_p(\mathbb{R}^n)$, $g \in L_q(\mathbb{R}^n)$ and $h \in L_r(\mathbb{R}^n)$. Then

$$||f * g * h||_{\infty} \le (C_p C_r C_r)^n ||f||_p ||g||_q ||h||_r$$

This inequality easily tensorizes. To see this suppose that such an inequality is true on \mathbb{R}^n with constant C(n) and on \mathbb{R}^n with constant C(m). Then on \mathbb{R}^{n+m} we can write

$$(f * g * h)(x_1, x_2) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} f(x_1 - y_1 - z_1, x_2 - y_2 - z_2) g(y_1, y_2) h(z_1, z_2) dy_2 dz_2 dy_1 dz_1$$

$$\leq C(m) \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} f(x_1 - y_1 - z_1, t)^p dt \right)^{1/p} \left(\int_{\mathbb{R}^m} g(y_1, t)^q dt \right)^{1/q} \left(\int_{\mathbb{R}^m} h(z_1, t)^r dt \right)^{1/r} dy_1 dz_1$$

$$\leq C(n) C(m) \|f\|_p \|g\|_q \|h\|_r.$$

Thus if the inequality holds true on \mathbb{R} with some constant C(1) then $C(n) \leq C(1)^n$. To see that in fact the optimal constants satisfy $C(n) = C(1)^n$ it suffices to take product functions.

2. Entropy power inequality

Let f be a density of a random vector X having values in \mathbb{R}^n . Take p > 0. We define the Rényi entropy via

$$h_p(X) = \frac{1}{1-p} \ln\left(\int f^p\right) = \frac{p}{1-p} \ln \|f\|_p.$$

Here we will be always assuming that the integrals are finite. We also define

$$h(X) = \lim_{p \to 1} h_p(X) = -\int f \ln f.$$

Suppose A is an invertible map. We claim that

(14)
$$h_p(AX) = h_p(X) + \ln |\det(A)|$$

Indeed, the density of AX is $f_{AX}(x) = \frac{1}{|\det(A)|} f(A^{-1}x)$. Thus

$$h_p(AX) = \frac{1}{1-p} \ln \int f_{AX}^p = \frac{1}{1-p} \ln \int \left(\frac{1}{|\det(A)|} f(A^{-1}x)\right)^p dx$$

= $\frac{1}{1-p} \ln \int \left(\frac{1}{|\det(A)|} f(y)\right)^p |\det(A)| dy = \frac{1}{1-p} \ln \int f + \ln |\det A|$
= $h_p(X) + \ln |\det A|.$

Suppose f is the density of X and g is the density of Y with X, Y independent. Suppose p, q, r > 1 with $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$. Then taking the logarithm of 13 we get

$$\ln \|f * g\|_{r} \le n \ln \left(\frac{C_{p}C_{q}}{C_{r}}\right) + \ln \|f\|_{p} + \|g\|_{q}$$

Note that f * g is the density of X + Y. Thus the above is equivalent to

$$\frac{1-r}{r}h_r(X+Y) \le n\ln\left(\frac{C_pC_q}{C_r}\right) + \frac{1-p}{p}h_p(X) + \frac{1-q}{q}h_q(Y).$$

This is (remember that r > 1)

(15)
$$h_r(X+Y) \ge -\frac{r}{r-1}n\ln\left(\frac{C_pC_q}{C_r}\right) + \frac{r}{r-1}\cdot\frac{p-1}{p}h_p(X) + \frac{r}{r-1}\cdot\frac{q-1}{q}h_q(Y).$$

Let us fix $\lambda \in (0, 1)$ and r > 1. Define

$$p = \frac{1}{1 - \lambda + \frac{\lambda}{r}}, \qquad q = \frac{1}{\lambda + \frac{1 - \lambda}{r}}.$$

These numbers clearly satisfy $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$. We have

$$\frac{r}{r-1} \cdot \frac{p-1}{p} = \lambda \qquad \frac{r}{r-1} \cdot \frac{q-1}{q} = 1 - \lambda.$$

We got

$$h_r(X+Y) \ge -\frac{r}{r-1}n\ln\left(\frac{C_pC_q}{C_r}\right) + \lambda h_p(X) + (1-\lambda)h_q(Y).$$

Using the scaling (14) and taking $\sqrt{\lambda}X$ and $\sqrt{1-\lambda}Y$ instead of X and Y, we obtain

$$h_r(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) \ge -\frac{r}{r-1}n\ln\left(\frac{C_pC_q}{C_r}\right) + \lambda h_p(\sqrt{\lambda}X) + (1-\lambda)h_q(\sqrt{1-\lambda}Y)$$
$$= -\frac{r}{r-1}n\ln\left(\frac{C_pC_q}{C_r}\right) + \lambda h_p(X) + (1-\lambda)h_q(Y)$$
$$+ \frac{n}{2}\left(\lambda\ln\lambda + (1-\lambda)\ln(1-\lambda)\right).$$

Rewriting gives

$$h_r(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) - \lambda h_p(X) - (1-\lambda)h_q(Y) \ge -\frac{r}{r-1}n\ln\left(\frac{C_pC_q}{C_r}\right) + \frac{n}{2}\left(\lambda\ln\lambda + (1-\lambda)\ln(1-\lambda)\right)$$

Remember that $p = p(\lambda, r)$ and $q = q(\lambda, r)$. If λ is fixed and $r \to 1^+$ we get $p \to 1^+$ and $q \to 1^+$. The left hand side converges to

$$h(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) - \lambda h(X) - (1-\lambda)h(Y).$$

We claim that the right hand side converges to 0. It suffices to verify that

$$\lim_{r \to 1^+} \frac{r}{r-1} \ln\left(\frac{C_p C_q}{C_r}\right) = \frac{1}{2} \left(\lambda \ln \lambda + (1-\lambda) \ln(1-\lambda)\right).$$

In other words, using the definition of C_p, C_q and C_r , we have to show that

$$\lim_{r \to 1^+} \frac{r}{r-1} \ln\left(\frac{p^{1/p} q^{1/q}}{p'^{1/p'} q'^{1/q'}} \cdot \frac{r'^{1/r'}}{r^{1/r}}\right) = \lim_{r \to 1^+} \frac{r}{r-1} \ln\left(\frac{p^{1/p} q^{1/q}}{r^{1/r}}\right) + \lim_{r \to 1^+} \frac{r}{r-1} \ln\left(\frac{r'^{1/r'}}{p'^{1/p'} q'^{1/q'}}\right).$$

Since $p' = r'/\lambda$ and $q' = r'/(1-\lambda)$, we have

Since $p' = r'/\lambda$ and $q' = r'/(1 - \lambda)$, we have

$$\frac{r}{r-1}\ln\left(\frac{r'^{1/r'}}{p'^{1/p'}q'^{1/q'}}\right) = \frac{r}{r-1}\ln\left(\frac{r'^{1/r'}}{\left(\frac{r'}{\lambda}\right)^{\frac{\lambda}{r'}}\left(\frac{r'}{1-\lambda}\right)^{\frac{1-\lambda}{r'}}}\right) = \frac{r}{r-1}\cdot\frac{1}{r'}\left(\lambda\ln\lambda + (1-\lambda)\ln(1-\lambda)\right)$$
$$= \lambda\ln\lambda + (1-\lambda)\ln(1-\lambda).$$

We also have

$$\frac{r}{r-1}\ln\left(\frac{p^{1/p}q^{1/q}}{r^{1/r}}\right) = \frac{r}{r-1}\left(\frac{1}{p}\ln p + \frac{1}{q}\ln q\right) - \frac{\ln r}{r-1}$$

We have $\frac{\ln r}{r-1} \to 1$ when $r \to 1^+$. Moreover,

$$\frac{r}{1-r}\left(\left(1-\lambda+\frac{\lambda}{r}\right)\ln\left(1-\lambda+\frac{\lambda}{r}\right)+\left(\lambda+\frac{1-\lambda}{r}\right)\ln\left(\lambda+\frac{1-\lambda}{r}\right)\right)$$
$$\sim_{r\to 1^{+}}\frac{r}{1-r}\left(\left(1-\lambda+\frac{\lambda}{r}\right)\left(-\lambda+\frac{\lambda}{r}\right)+\left(\lambda+\frac{1-\lambda}{r}\right)\left(-1+\lambda+\frac{1-\lambda}{r}\right)\right)$$
$$=\left(1-\lambda+\frac{\lambda}{r}\right)\cdot\lambda+\left(\lambda+\frac{1-\lambda}{r}\right)\cdot\left(1-\lambda\right)\xrightarrow[r\to 1^{+}]{}\lambda+\left(1-\lambda\right)=1.$$

The claim is now established and we arrive at

$$h(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) \ge \lambda h(X) + (1-\lambda)h(Y).$$

This is the linear form of the so-called entropy power inequality. Let us define

$$N(X) = \frac{1}{2\pi e} \exp\left(\frac{2h(X)}{n}\right).$$

Suppose X is a Gaussian vector with the covariance matrix K_X . Then the density of this vector is equal to

$$\varphi_G(x) = \frac{1}{2\pi\sqrt{\det K_G}}\varphi_n(|K_G^{-1/2}|).$$

In other words, $X = (K_X)^{1/2}G$, where $G \sim \mathcal{N}(0, I)$. Thus $h(X) = h(G) + \frac{1}{2} \ln \det K_X$ and

$$N(X) = \det(K_X)^{1/n} N(G) = \det(K_X)^{1/n}$$

since N(G) = 1 as

$$h(G) = -\int \varphi_n \ln \varphi_n = -\int \varphi_n \ln \left((2\pi)^{-n/2} \exp(-|x|^2/2) \right) = \frac{n}{2} \ln(2\pi) + \frac{n}{2}$$

We give three equivalent formulations of the entropy power inequality.

Theorem 32. Let X, Y be independent random variables having values is \mathbb{R}^n . We have

(i) We have

 $N(X+Y) \ge N(X) + N(Y)$

or equivalently

$$e^{\frac{2}{n}h(X+Y)} \ge e^{\frac{2}{n}h(X)} + e^{\frac{2}{n}h(Y)}.$$

(ii) For any $\lambda \in [0, 1]$ we have

$$h(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) \ge \lambda h(X) + (1-\lambda)h(Y).$$

(iii) If G_X and G_Y are independent multiples of a standard Gaussian such that $h(G_X) = h(X)$ and $h(G_Y) = h(Y)$ then

$$N(X+Y) \ge N(G_X+G_Y)$$
 or $h(X+Y) \ge h(G_X+G_Y)$.

Proof. To show that (iii) implies (ii) we observe that if K_X and K_Y are the covariance matrices of G_X and G_Y then

$$h(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) \ge h(\sqrt{\lambda}G_X + \sqrt{1-\lambda}G_Y) \ge \lambda h(G_X) + (1-\lambda)h_G(Y),$$

where the last inequality is equivalent to

$$N(\sqrt{\lambda}G_X + \sqrt{1 - \lambda}G_Y) \ge N(G_X)^{\lambda}N(G_Y)^{1 - \lambda}$$

follows from

$$\det(\lambda G_X + (1-\lambda)G_Y) \ge \det(G_X)^{\lambda} \det(G_Y)^{1-\lambda}$$

We now show that (ii) implies (i). Replacing X with X/λ and Y with $Y/(1-\lambda)$ in (ii) gives

$$h(X+Y) \ge \lambda h(X) + (1-\lambda)h(Y) - \frac{n}{2}\left(\lambda \ln \lambda + (1-\lambda)\ln(1-\lambda)\right)$$

We shall optimize the right hand side with respect to $\lambda \in [0, 1]$. Computing the derivative gives

(16)
$$h(X) - h(Y) = \frac{n}{2} \ln\left(\frac{\lambda}{1-\lambda}\right),$$

which yields

$$\lambda = \frac{e^{\frac{2}{n}h(X)}}{e^{\frac{2}{n}h(X)} + e^{\frac{2}{n}h(Y)}}.$$

Note that (16) is equivalent to $h(X) - \frac{n}{2} \ln \lambda = h(Y) - \frac{n}{2} \ln(1-\lambda)$. Thus for optimal lambda we have

$$\begin{split} h(X+Y) &\geq \lambda h(X) + (1-\lambda)h(Y) - \frac{n}{2}\left(\lambda \ln \lambda + (1-\lambda)\ln(1-\lambda)\right) \\ &= \lambda \left(h(X) - \frac{n}{2}\ln \lambda\right) + (1-\lambda)\left(h(Y) - \frac{n}{2}\ln(1-\lambda)\right) \\ &= h(X) - \frac{n}{2}\ln \lambda = \frac{n}{2}\ln\left(e^{\frac{2}{n}h(X)} + e^{\frac{2}{n}h(Y)}\right). \end{split}$$

Rearranging gives (i).

To show that (i) implies (iii) we observe that

$$N(X+Y) \ge N(X) + N(Y) = N(G_X) + N(G_Y) = \det(K_X)^{1/n} + \det(K_Y)^{1/n}$$

= det(K_X + K_Y)^{1/n} = N(G_X + G_Y),

where the third equality follows from the fact that K_X and K_Y are multiples of the identity matrix.

2.1. Geometric Brascamp-Lieb inequality. There is a setting of Brascamp-Lieb type inequality where the optimal constant is D = 1.

Theorem 33. Let $n, m \ge 1$ and let $u_1, \ldots, u_m \in S^{n-1}, c_1, \ldots, c_m > 0$ be such that $I = \sum_{j=1}^m c_j u_j \otimes u_j$. If $f_1, \ldots, f_m : \mathbb{R} \to \mathbb{R}_+$ are integrable functions then

(17)
$$\int_{\mathbb{R}^n} \prod_{j=1}^m \left(f_j(\langle x, u_j \rangle) \right)^{c_j} \, \mathrm{d}x \le \prod_{j=1}^m \left(\int_{\mathbb{R}} f_j \right)^{c_j}$$

Remark 4. The condition on c_i 's in the Brascamp-Lieb inequality is satisfied in the above setting. Indeed, we have $n_1 = \ldots n_m = 1$ and

$$n = \operatorname{tr}(\mathbf{I}) = \sum_{j=1}^{m} c_j \operatorname{tr}(u_j \otimes u_j) = \sum_{j=1}^{m} c_j |u_j|_2^2 = \sum_{j=1}^{m} c_j = \sum_{j=1}^{m} c_j n_j.$$

Remark 5. The condition

(18)
$$\mathbf{I} = \sum_{j=1}^{m} c_j u_j \otimes u_j$$

is equivalent to

$$\forall x \in \mathbb{R}^n, \quad x = \sum_{j=1}^m c_j \langle x, u_j \rangle u_j \quad \text{or equivalently to} \quad \forall x \in \mathbb{R}^n, \quad |x|_2^2 = \sum_{j=1}^m c_j \langle x, u_j \rangle^2.$$

We can easily construct examples of vectors satisfying condition (18). Let H be an ndimensional subspace of \mathbb{R}^m . Let e_1, \ldots, e_m be the standard orthonormal basis in \mathbb{R}^m and let $P : \mathbb{R}^m \to H$ be the orthogonal projection onto H. Clearly, $I_{\mathbb{R}^m} = \sum_{j=1}^m e_j \otimes e_j$ and $x = \sum_{j=1}^m \langle x, e_j \rangle e_j$, hence $Px = \sum_{j=1}^m \langle x, e_j \rangle Pe_j$. If $x \in H$ then Px = x and $\langle x, e_j \rangle =$ $\langle Px, e_j \rangle = \langle x, Pe_j \rangle$, therefore $x = \sum_{j=1}^m \langle x, Pe_j \rangle Pe_j$. Thus $I_{H \approx \mathbb{R}^n} = \sum_{j=1}^m c_j u_j \otimes u_j$, where $c_j = |Pe_j|^2$ and $u_j = Pe_j/|Pe_j|$.

To prove Theorem 33 it suffice to prove that D = 1 in the Brascamp-Lieb setting. Namely, to show that the condition

$$\sum_{i=1}^{m} v_i \otimes v_i = I_n$$

for some v_1, \ldots, v_m in \mathbb{R}^n implies

$$\det\left(\sum_{i=1}^m a_i v_i \otimes v_i\right) \ge \prod_{i=1}^m a_i^{|v_i|^2}, \qquad a_1, \dots, a_m > 0.$$

We then use it with $v_i = \sqrt{c_i}u_i$ to verify that the optimal constant D in the formulation of BL is equal 1.

Using Cauchy-Binet formula we have for any $n \times m$ matrix A and $m \times n$ matrix B that

$$\det(AB) = \sum_{|I|=n} \det(A_I) \det(B^I),$$

where the sums runs over all subsets $I \subseteq \{1, \ldots, m\}$ of cardinality n and A_I is obtained by keeping only the columns indexed by elements of I, whereas B^I is obtained by keeping only the rows indexed by elements of I. Since

$$\sum_{i=1}^{m} a_i v_i \otimes v_i = \begin{bmatrix} | & | \\ \sqrt{a_1} v_1 & \dots & \sqrt{a_m} v_m \\ | & | \end{bmatrix} \cdot \begin{bmatrix} - & \sqrt{a_1} v_1 & - \\ \vdots \\ - & \sqrt{a_m} v_m & - \end{bmatrix}$$

we get

$$\det\left(\sum_{i=1}^{m} a_i v_i \otimes v_i\right) = \sum_{|I|=n} a_I d_I$$

where

$$d_I = (\det((v_i)_{i \in I}))^2$$
 and $a_I = \prod_{i \in I} a_i$.

Taking $a_1 = \ldots = a_m = 1$ gives $\sum_{|I|=n} d_I = \det(\sum v_i \otimes v_i) = \det(I) = 1$. Thus, by AM-GM we get

$$\sum_{|I|=n} a_I d_I \ge \prod_{|I|=n} a_I^{d_I} = \prod_{i=1}^m a_i^{\sum_{I:i \in I, |I|=n} d_I}.$$

But

$$\sum_{I:i\in I} d_I = \sum_{|I|=n} d_I - \sum_{I:i\notin I, |I|=n} d_I = 1 - \det\left(\sum_{k\neq i} v_k \otimes v_k\right)$$
$$= 1 - \det(I_n - v_i \otimes v_i) = 1 - \det(I_n - v_i v_i^T) = 1 - \det(I_1 - v_i^T v_i)$$
$$= 1 - (1 - v_i^T v_i) = |v_i|^2.$$

Note that we have used the Sylvester identity.

Lemma 13. Suppose X is a $m \times n$ matrix and Y is a $n \times m$ matrix. Then

$$\det(I_m + XY) = \det(I_n + YX).$$

Proof. To prove this, let us first observe that we have the identity

$$\left(\begin{array}{cc}I_n & -Y\\X & I_m\end{array}\right) \cdot \left(\begin{array}{cc}I_n & Y\\0 & I_m\end{array}\right) = \left(\begin{array}{cc}I_n & 0\\X & XY + I_m\end{array}\right)$$

We have

$$\det\left(\left(\begin{array}{cc}I_n & -Y\\X & I_m\end{array}\right)\cdot\left(\begin{array}{cc}I_n & Y\\0 & I_m\end{array}\right)\right) = \det\left(\begin{array}{cc}I_n & 0\\X & XY+I_m\end{array}\right) = \det(XY+I_m)$$

Since det(AB) = det(BA), the left hand side is the same as

$$\det\left(\left(\begin{array}{cc}I_n & Y\\0 & I_m\end{array}\right)\cdot\left(\begin{array}{cc}I_n & -Y\\X & I_m\end{array}\right)\right) = \det\left(\begin{array}{cc}I_n + YX & 0\\X & I_m\end{array}\right) = \det(I_n + YX).$$

2.2. John's ellipsoid theorem. We shall prove the following classical fact.

Theorem 34. Let K be a convex body (compact convex set with non-empty interior). Then

- (i) There exists a unique ellipsoid $\mathcal{E}_K \subseteq K$ with maximal volume.
- (ii) If $B_2^n \subset K$ is the ellipsoid of maximal volume contained in a symmetric convex body $K \subset \mathbb{R}^n$ then there exist $c_1, \ldots, c_m > 0$ and contact points $u_1, \ldots, u_m \in \mathbb{R}^n$ such that $|u_j|_2 = ||u_j||_K = ||u_j||_{K^\circ} = 1$ for $1 \leq j \leq m$ and

(19)
$$I_{\mathbb{R}^n} = \sum_{j=1}^m c_j u_j \otimes u_j.$$

- (iii) For symmetric convex body K we have $\mathcal{E}_K \subseteq K \subseteq \sqrt{n}\mathcal{E}_K$.
- (iv) If $B_2^n \subseteq K$ and there exist contact points u_1, \ldots, u_m of B_2^n and K, and numbers c_1, \ldots, c_m such that (19) is satisfied, then $\mathcal{E}_K = B_2^n$.

Remark 6. Let K be symmetric. One can define contact points of B_2^n with K as points satisfying $u \in \partial K \cap S^{n-1}$. In other words, this means $|u| = ||u||_K = 1$. We claim that in this situation also $||u||_{K^\circ} = 1$. Let us take a contact point u and let H be a supporting hyperplane of K at u. We claim that $H = \{x : \langle x, u \rangle = 1\}$. From the inclusion $B_2^n \subseteq K$ and from the fact that $u \in \partial K \cap \partial B_2^n$ we see that H is also a supporting hyperplane of B_2^n at u. Thus H is unique and equal to $\{x : \langle x, u \rangle = 1\}$, which is clearly the supporting hyperplane for the ball B_2^n . It follows that $K \subseteq \{x : \langle x, u \rangle \leq 1\}$ and by symmetry of K we get $K \subseteq \{x : |\langle x, u \rangle | \leq 1\}$. Since this is true for any contact point, we in fact get

$$K \subseteq \bigcap_{i=1}^{m} \{ x : |\langle x, u_i \rangle| \le 1 \},\$$

which we shall use in the sequel.

For any u in \mathbb{R}^n we have

$$\|u\|_{K^{\circ}} = \sup\left\{ |\langle x, u \rangle| : x \in K \right\}.$$

If u is a contact point then by the fact that $K \subseteq \{x : |\langle x, u \rangle| \leq 1\}$ we get $||u||_{K^{\circ}} \leq 1$. In fact we have equality by taking $x = u \in K$ and using the fact that |u| = 1.

Proof. (i) By applying a suitable linear transformation we can assume that B_2^n is the ellipsoid of maximal volume. Suppose there is some other ellipsoid \mathcal{E} contained in K and having the same volume. There is an invertible linear transformation T and a vector x_0 such that $\mathcal{E} = TB_2^n + x_0$. Note that

$$\mathcal{E} - x_0 = T\{\langle x, x \rangle \le 1\} = \{x : \langle T^{-1}x, T^{-1}x \rangle \le 1\} = \{x : \langle T^{-1}x, T^{-1}x \rangle \le 1\}$$
$$= \{x : \langle x, (T^{-1})^*T^{-1}x \rangle \le 1\} = \{x : \langle ((T^{-1})^*T^{-1})^{1/2}x, ((T^{-1})^*T^{-1})^{1/2}x \rangle \le 1\}$$
$$= (((T^{-1})^*T^{-1})^{1/2})^{-1}B_2^n = (TT^*)^{1/2}B_2^n,$$

where we have used the fact that a positive definite map $(T^{-1})^*T^{-1}$ has a square root. Thus, we can assume that the map T is positive definite.

Consider the ellipsoid

$$\mathcal{E}_0 = \frac{x_0}{2} + \frac{I+T}{2}B_2^n \subseteq \frac{B_2^n + TB_2^n + x_0}{2} \subseteq K,$$

where the last inclusion follows by the convexity of K and the fact that both B_2^n and \mathcal{E} are subsets of K. We shall show that the volume of \mathcal{E}_0 is strictly bigger that the volume of B_2^n . This will be a contradiction. Note that

$$|\mathcal{E}_0| = \det\left(\frac{I+T}{2}\right)|B_2^n| \ge \sqrt{\det(A)}|B_2^n| = |B_2^n|,$$

since $|B_2^n| = |\mathcal{E}| = \det(T)|B_2^n|$ and thus $\det(T) = 1$. The above inequality is the concavity of the determinant proved earlier. By maximality of B_2^n we see that there has to be equality in the above bound, which gives A = I. Thus, $\mathcal{E} = B_2^n + x_0$. we have $\mathcal{E}_1 := B_2^n + \frac{x_0}{2} = \frac{1}{2}B_2^n + \frac{1}{2}\mathcal{E} \subseteq K$. Since $\operatorname{conv}(B_2^n, \mathcal{E}) \subseteq K$, it is easy to see that one can dilate \mathcal{E}_1 a bit in the direction of $[0, x_0]$ to get a bigger ellipsoid contained in K. This is a contradiction.

(ii) Step 1. Since (19) implies that $\sum_{i=1}^{m} c_i = n$, we have to show that $\frac{I_n}{n} \in \text{conv}(C)$, where

$$C = \{ u \otimes u : \ |u| = \|u\|_K = 1 \}.$$

Assume by contradiction that it is not possible. If we view C as a subset of the space \mathbb{R}^{n^2} , the set $\operatorname{conv}(C)$ is compact and convex. Thus, if $\frac{I_n}{n} \notin \operatorname{conv} C$, we can find a functional ϕ (viewed as a $n \times n$ matrix) and a real number r such that

(20)
$$\left\langle \phi, \frac{I_n}{n} \right\rangle < r < \left\langle \phi, u \otimes u \right\rangle$$

for all u such that $|u| = ||u||_{K}$. Here $\langle (a_{ij})_{i,j=1}^{n}, (b_{ij})_{i,j=1}^{n} \rangle = \sum_{i,j=1}^{n} a_{ij} b_{ij}$.

Step 2. We can assume that ϕ is a symmetric matrix. Indeed, for symmetric A we have $\langle \phi, A \rangle = \langle \phi^*, A \rangle$, where ϕ^* is the transpose of ϕ . Thus, if ϕ is not symmetric, we can replace ϕ with $\frac{\phi + \phi^*}{2}$, not changing (20).

Step 3. Since $\operatorname{tr}(\frac{I_n}{n}) = 1 = \operatorname{tr}(u \otimes u)$, we can add cI to the matrix ϕ , not changing the separation property (take s = r + c instead of r). Thus, we can assume that $\operatorname{tr}(\phi) = 0$, which means that $\phi(I_n) = 0$. Thus, we found a functional matrix B and a real number s such that for all contact point

$$\langle B, u \otimes u \rangle > s > 0$$
 and $\operatorname{tr}(B) = 0.$

Note that

$$0 < s < \langle B, u \otimes u \rangle = \sum_{j,k=1}^{n} B_{jk} (u \otimes u)_{jk} = \sum_{j,k=1}^{n} B_{jk} u_j u_k = u^* B u = \langle Bu, u \rangle.$$

Step 4. Define

$$\mathcal{E}_{\delta} = \{x \in \mathbb{R}^n : \langle (I_n + \delta B)x, x \rangle \le 1\}.$$

For small $\delta > 0$ this is an ellipsoid approaching B_2^n when $\delta \to 0^+$. We shall show that for small $\delta > 0$ this ellipsoid is contained in K and $|\mathcal{E}_{\delta}| > |B_2^n|$.

Step 5. We verify that indeed $|\mathcal{E}_{\delta}| > |B_2^n|$. For an invertible symmetric matrix A the set $\{\langle Ax, x \rangle \leq 1\}$ is an image of B_2^n under $A^{-1/2}$. Indeed

$$\{\langle Ax, x \rangle \le 1\} = \{\langle A^{1/2}x, A^{1/2}x \rangle \le 1\} = A^{-1/2}B_2^n.$$

Thus

$$|\mathcal{E}_{\delta}| = \frac{|B_2^n|}{\det(I_n + \delta B)^{1/2}}.$$

However, by AM-GM we have

$$\det(I_n + \delta B)^{1/n} < \frac{\operatorname{tr}(I_n + \delta B)}{n} = 1,$$

since tr(B) = 0. The inequality is strict since not all the eigenvalues of B are equal (otherwise tr(B) = 0 would imply that all the eigenvalue are zero, which would mean that B = 0).

Step 6. Let U be the set of contact points. We define

$$S_{+} = \left\{ v \in S^{n-1} : \operatorname{dist}(v, U) \le \frac{s}{4\|B\|} \right\}, \qquad S_{-} = \left\{ v \in S^{n-1} : \operatorname{dist}(v, U) \ge \frac{s}{4\|B\|} \right\}.$$

Clearly $S_+ \cup S_- = S^{n-1}$. Our goal is to show that (for small $\delta > 0$) $\partial K \ni v / ||v||_K \notin \mathcal{E}_{\delta}$ for all unit vector v, which easily implies that $\mathcal{E}_{\delta} \subseteq K$.

Step 7. We first check it for $v \in S_-$. By compactness there is $\varepsilon > 0$ such that $dist(\partial K, S_-) \ge \varepsilon > 0$ and thus

dist
$$\left(\left\{\frac{v}{\|v\|_{K}}, v \in S_{-}\right\}, B_{2}^{n}\right) \ge \varepsilon > 0.$$

Thus

dist
$$\left(\left\{\frac{v}{\|v\|_{K}}, v \in S_{-}\right\}, \left(1 + \frac{\varepsilon}{2}\right)B_{2}^{n}\right) \geq \frac{\varepsilon}{2} > 0$$

The assertion follows by observing that for sufficiently small $\delta > 0$ we have $\mathcal{E}_{\delta} \subseteq (1 + \frac{\varepsilon}{2}) B_2^n$.

Step 8. The case $v \in S_+$ is more delicate. By Step 3, for every $u \in U$ we have

(21)
$$\langle (I_n + \delta B)u, u \rangle \ge 1 + \delta s.$$

Furthermore, if $v \in S_+$ then

$$\begin{aligned} |\langle (I_n + \delta B)v, v \rangle - \langle (I_n + \delta B)u, u \rangle| &= \delta |\langle Bv, v \rangle - \langle Bu, u \rangle| \\ &\leq \delta |\langle Bv, v \rangle - \langle Bv, u \rangle| + \delta |\langle Bv, u \rangle - \langle Bu, u \rangle| \\ &\leq 2\delta ||B|| |u - v| \leq \frac{1}{2} s \delta. \end{aligned}$$

This together with (21) yields $\langle (I_n + \delta B)v, v \rangle \geq 1 + \frac{1}{2}\delta s > 1$ and thus $v \notin \mathcal{E}_{\delta}$. Since $v \in B_2^n \subseteq K$, we have $\|v\|_K \leq 1$ and thus also $v/\|v\|_K \notin \mathcal{E}_{\delta}$.

This finishes the proof of point (ii).

(iii) Without loss of generality, by applying linear transformation we can assume that $\mathcal{E}_K = B_2^n$. Then

$$K \subseteq \bigcap_{i=1}^{m} \{ |\langle x, u_i \rangle| \le 1 \}.$$

Thus, if $x \in K$ then

$$|x|^2 = \sum_{i=1}^m c_i \langle x, u_i \rangle^2 \le \sum_{i=1}^m c_i = n.$$

Thus $|x| \leq \sqrt{n}$, which means that $x \in \sqrt{n}B_2^n$.

(iv) Clearly by uniqueness of maximal ellipsoid \mathcal{E}_K is symmetric if the body is symmetric. Take an ellipsoid

$$\mathcal{E} = \{ x : \sum_{i=1}^{n} \frac{\langle x, e_j \rangle^2}{\alpha_j^2} \le 1 \},\$$

where (e_j) is some orthonormal basis. Suppose $u \in \{u_1, \ldots, u_m\}$ be a contact point. Take a point $y = \sum_{j=1}^n \alpha_j \langle u, e_j \rangle e_j$. We claim that $y \in \mathcal{E}$. Indeed, $\langle y, e_j \rangle = \alpha_j \langle u, e_j \rangle$ and thus

$$\sum_{i=1}^{n} \frac{\langle y, e_j \rangle^2}{\alpha_j^2} = \sum_{i=1}^{n} \langle u, e_j \rangle^2 = |u|^2 = 1.$$

Since

$$\sum_{j=1}^{n} \alpha_j \langle u, e_j \rangle e_j = y \in \mathcal{E} \subseteq K \subseteq \bigcap_{i=1}^{m} \{x : |\langle x, u_i \rangle| \le 1\} \subseteq \{x : |\langle x, u \rangle| \le 1\},\$$

we get

$$\left\langle \sum_{j=1}^{n} \alpha_j \left\langle u, e_j \right\rangle e_j, u \right\rangle = \sum_{j=1}^{n} \alpha_j \left\langle u, e_j \right\rangle^2 \le 1.$$

But, using (19), this means that

$$\sum_{j=1}^{n} \alpha_j^2 = \sum_{j=1}^{n} \alpha_j^2 |e_j|^2 = \sum_{j=1}^{n} \alpha_j^2 \sum_{i=1}^{m} c_i \langle u_i, e_j \rangle^2 = \sum_{i=1}^{m} c_i \sum_{j=1}^{n} \alpha_j^2 \langle u_i, e_j \rangle^2 \le \sum_{i=1}^{m} c_i = n.$$

Thus

$$\left(\frac{|\mathcal{E}|}{|B_2^n|}\right)^{2/n} = \left(\prod_{i=1}^n \alpha_i^2\right)^{1/n} \le \frac{1}{n} \sum_{j=1}^n \alpha_j^2 \le 1.$$

Example 3. If $K = B_{\infty}^{n}$ then $\mathcal{E}_{K} = B_{2}^{n}$. Indeed, it follows from point (iv) of John's theorem as $\pm e_{i}$ are contact points and $I = \frac{1}{2} \sum_{i=1}^{n} e_{i} \otimes e_{i} + \frac{1}{2} \sum_{i=1}^{n} (-e_{i}) \otimes (-e_{i})$.

2.3. Reverse isoperimetric inequality. Let us state the reverse isoperimetric inequality.

Theorem 35. Let K be a symmetric convex body in \mathbb{R}^n . Then there exists an affine transformation \widetilde{K} of K such that

(22)
$$|\widetilde{K}| = |B_{\infty}^{n}|, \text{ and } |\partial\widetilde{K}| \le |\partial B_{\infty}^{n}|$$

or equivalently

(23)
$$\frac{|\partial K|}{|K|^{\frac{n-1}{n}}} \le \frac{|\partial B_{\infty}^{n}|}{|B_{\infty}^{n}|^{\frac{n-1}{n}}} = 2n$$

Before we give a proof of Theorem 35 we introduce the notion of the volume ratio.

Definition 1. Let $K \subset \mathbb{R}^n$ be a convex body. The volume ratio of K is defined as

$$vr(K) = \inf\left\{\left(\frac{|K|}{|\mathcal{E}|}\right)^{1/n}, \quad \mathcal{E} \subset K \text{ is an ellipsoid}\right\}.$$

The ellipsoid of maximal volume contained in K is called the John ellipsoid. If the John ellipsoid of K is equal to B_2^n then we say that K is in the John position.

We have the following theorem.

Theorem 36. For every symmetric convex body $K \subset \mathbb{R}^n$ we have

(24)
$$vr(K) \le vr(B_{\infty}^{n}) = \frac{2}{(|B_{2}^{n}|)^{1/n}}.$$

Theorem $34 \implies$ Theorem 36. The quantity vr(K) is invariant under invertible linear transformations. We let as an exercise to check that the ellipsoid of maximal volume contained in K is unique. Therefore we may assume that the John ellipsoid of K is B_2^n . Using Theorem 34 we find numbers $c_1, \ldots, c_m > 0$ and unit vectors $u_1, \ldots, u_m \in \mathbb{R}^n$ on the boundary of Ksuch that

$$\mathbf{I}_{\mathbb{R}^n} = \sum_{j=1}^m c_j u_j \otimes u_j.$$

Since $u_j \in \partial B_2^n \cap \partial K$ and K is symmetric we get

$$K \subset K' := \{x \in \mathbb{R}^n, |\langle x, u_j \rangle| \le 1, \text{ for all } 1 \le j \le m\}.$$

Let $f_j(t) = \mathbf{1}_{[-1,1]}(t)$ for $1 \leq j \leq m$. Note that $f_j = f_j^{c_j}$, $1 \leq j \leq m$. From Theorem 33 we have

$$|K| \le |K'| = \int_{\mathbb{R}^n} \prod_{j=1}^m f_j^{c_j} \left(\langle x, u_j \rangle \right) \, \mathrm{d}x \le \prod_{j=1}^m \left(\int f_j \right)^{c_j} = 2^{\sum_{j=1}^m c_j} = 2^n = |B_\infty^n|.$$

From Example 3 we know that B_2^n is the John ellipsoid for the cube B_{∞}^n . Therefore

$$vr(B_{\infty}^{n}) = \frac{2}{(|B_{2}^{n}|)^{1/n}}.$$

We now show that Theorem 36 implies Theorem 35.

Proof of Theorem 35. Let \widetilde{K} be the linear image of K such that $B_2^n \subset \widetilde{K}$ is the John ellipsoid of \widetilde{K} . By Theorem 36 we have $|\widetilde{K}| \leq 2^n$. Hence,

$$\begin{split} |\partial \widetilde{K}| &= \liminf_{\varepsilon \to 0^+} \frac{|\widetilde{K} + \varepsilon B_2^n| - |\widetilde{K}|}{\varepsilon} \le \liminf_{\varepsilon \to 0^+} \frac{|\widetilde{K} + \varepsilon \widetilde{K}| - |\widetilde{K}|}{\varepsilon} \\ &= n|\widetilde{K}| = n|\widetilde{K}|^{\frac{n-1}{n}} \cdot |\widetilde{K}|^{\frac{1}{n}} \le 2n|\widetilde{K}|^{\frac{n-1}{n}}. \end{split}$$

This finishes the proof as the ratio $\frac{|\partial K|}{|K|^{\frac{n-1}{n}}}$ is affine invariant.

We state yet another application of the geometric Brascamp-Lieb inequality.

Theorem 37. If K is a symmetric convex body in the John position then $\mathbb{E} ||G||_K \geq \mathbb{E} |G|_{\infty}$, where G is the standard Gaussian vector in \mathbb{R}^n , i.e. the vector (g_1, \ldots, g_n) where $(g_i)_{i \leq n}$ are independent standard Gaussian random variables.

Proof. As in the proof of Theorem 35 we consider numbers $c_1, \ldots, c_m > 0$ and vectors u_1, \ldots, u_m satisfying the assertion of the Theorem 34. Note that

$$K \subset K' = \{ x \in \mathbb{R}^n, |\langle x, u_j \rangle | \le 1 \quad 1 \le j \le m \}.$$

Clearly,

$$||G||_{K} \ge ||G||_{K'} = \max_{1 \le j \le m} |\langle G, u_j \rangle|.$$

Moreover,

$$\mathbb{E} \left\| G \right\|_{K'} = \int_0^{+\infty} \mathbb{P} \left(\max_j \left| \langle G, u_j \rangle \right| \ge t \right) \, \mathrm{d}t.$$

We have $|G|_{\infty} = \max_{1 \le j \le m} |\langle G, e_j \rangle|$ so that

$$\mathbb{E}|G|_{\infty} = \int_{0}^{+\infty} \mathbb{P}\left(\max_{j} |\langle G, e_{j} \rangle| \ge t\right) dt = \int_{0}^{+\infty} \left(1 - \mathbb{P}\left(|g| \le t\right)^{n}\right) dt,$$

where g is the standard Gaussian random variable. To get the conclusion, it suffices to prove

$$\mathbb{P}\left(\max_{j} |\langle G, u_{j} \rangle| \le t\right) \le \left(\mathbb{P}\left(|g| \le t\right)\right)^{n}.$$

Take

$$h_j(s) = \mathbf{1}_{[-t,t]}(s) \frac{e^{-s^2/2}}{\sqrt{2\pi}}, \quad f_j(s) = \mathbf{1}_{[-t,t]}(s).$$

Since

$$|x|_2^2 = \sum_{j=1}^m c_j \langle x, u_j \rangle^2,$$

Theorem 33 implies that

$$\begin{split} \mathbb{P}\left(\max_{j} |\langle G, u_{j} \rangle| \leq t\right) &= \int_{\mathbb{R}^{n}} \mathbf{1}_{\{(\max_{j} |\langle x, u_{j} \rangle|) \leq t\}} \frac{1}{(2\pi)^{n/2}} e^{-|x|_{2}^{2}/2} \, \mathrm{d}x \\ &= \int_{\mathbb{R}^{n}} \prod_{j=1}^{m} f_{j}^{c_{j}} \left(\langle x, u_{j} \rangle\right) \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{|\langle x, u_{j} \rangle|^{2}}{2}\right)^{c_{j}} \, \mathrm{d}x \\ &= \int_{\mathbb{R}^{n}} \prod_{j=1}^{m} h_{j} \left(\langle x, u_{j} \rangle\right)^{c_{j}} \, \mathrm{d}x \\ &\leq \prod_{j=1}^{m} \left(\int h_{j}\right)^{c_{j}} = \left(\int_{-t}^{t} \frac{1}{\sqrt{2\pi}} e^{-u^{2}/2} \, \mathrm{d}u\right)^{n} \\ &= \left(\mathbb{P}\left(|g| \leq t\right)\right)^{n}, \end{split}$$

where we have used the fact that $\sum_{j=1}^{m} c_j = n$.

3. KLS LOCALIZATION

3.1. Topological vector spaces. Let us define topological vector spaces.

Definition 2. A vector space X (over \mathbb{R}) which is also equipped with some topology τ (family of open sets) is called a topological vector space if the singletons $\{x\}$ are closed sets and the operations $X \times X \to X$ given by $(x, y) \to x + y$ and $\mathbb{R} \times X \to X$ given by $(a, x) \to ax$ are continuous (the product spaces are equipped with product topologies and \mathbb{R} is equipped with the usual topology).

Note that it immediately follows that translations and multiplications by non-zero scalars are homeomorphisms of X. Thus, the topology τ is translation invariant – U is open if and only if x + U is open for ant $x \in X$.

Recall that a *neighbourhood* of $x \in X$ is any open set containing x. The collection $\tau' \subset \tau$ is a *base* for τ if every member of τ is a union of elements from τ' . A collection ω of neighbourhoods of $x \in X$ is a *local base* if every neighbourhood of x contains a member of ω . Note that by the translation invariance of τ the topology τ consists of all possible translates of neighbourhoods of 0.

Fact 38 (W. Rudin, Functional Analysis, 1.10). Suppose K is compact and C is closed in some t.v.s. and $K \cap C = \emptyset$. Then there is a neighbourhood V of 0 such that

$$(K+V) \cap (C+V) = \emptyset.$$

Proof. If W is a neighborhood of 0 then there is a symmetric neighborhood U of 0 such that $U + U \subseteq W$. Indeed, since the addition is continuous, there are neighborhood V_1, V_2 of 0 such that $V_1 + V_2 \subseteq W$. So it suffices to take $U = V_1 \cap V_2 \cap (-V_1) \cap (-V_2)$.

Applying the same trick for U we can get a symmetric neighborhood of 0 such that

$$U + U + U + U \subseteq W$$

and in particular $U + U + U \subseteq W$.

Now, we can assume that $K \neq \emptyset$. Take $x \in K$. We know that $x \notin C$. Using the translation invariance of the topology and the fact that $W = X \setminus C$ is a neighborhood of x, we get asymmetric neighborhood V_x of x such that $x + V_x + V_x + V_x \subseteq X \setminus C$ and thus $(x + V_x + V_x + V_x) \cap C = \emptyset$. By the symmetry of V_x we get $(x + V_x + V_x) \cap (C + V_x) = \emptyset$. By compactness one can find x_1, \ldots, x_n in K such that

$$K \subseteq (x_1 + V_{x_1}) \cup \ldots \cup (x_n + V_{x_n}).$$

Take $V = V_{x_1} \cap \ldots \cap V_{x_n}$. Then

$$K + V \subseteq \bigcup_{i=1}^{n} (x_i + V_{x_i} + V) \subseteq \bigcup_{i=1}^{n} (x_i + V_{x_i} + V_{x_i})$$

But for any i

$$(x_{i} + V_{x_{i}} + V_{x_{i}}) \cap (C + V) = (x_{i} + V_{x_{i}} + V_{x_{i}}) \cap (C + V_{x_{1}} \cap \ldots \cap V_{x_{n}})$$

= $(x_{i} + V_{x_{i}} + V_{x_{i}}) \cap (C + V_{x_{1}}) \cap \ldots \cap (C + V_{x_{n}})$
 $\subseteq (x_{i} + V_{x_{i}} + V_{x_{i}}) \cap (C + V_{x_{i}}) = \emptyset.$

Thus $(K+V) \cap (C+V) = \emptyset$.

Since C + V is a union of sets of the form c + V, $c \in C$, this set is open. Thus it is also true that

$$\overline{(K+V)} \cap (C+V) = \emptyset$$

In particular

 $\overline{(K+V)} \cap C = \emptyset.$

Taking $K = \{x\}$ (clearly they are compact by definition of compactness) we get the following fact.

Fact 39. Every neighbourhood of 0 contains a closure of some other neighbourhood of 0. In particular, every member of a local base at 0 contains a closure of some other member of a local base at 0.

Taking K and C to be singletons in the Fact 38 we get the following fact.

Fact 40. Every topological vector space is a Hausdorff space. In particular, every compact subset of a topological vector space is closed.

Fact 41. Let X be a t.v.s. and let $A, B \subseteq X$. Suppose λ is a real number. Then

(i)
$$\lambda \overline{A} = \overline{\lambda A}$$
,
(ii) $\overline{A} + \overline{B} \subseteq \overline{A + B}$,

(iii) if A is convex then \overline{A} is also convex.

Proof. (i) If $\lambda = 0$ this is obvious. If $\lambda \neq 0$ it follows from the fact that $f(x) = \lambda x$ is a homeomorphism as for homeomorphism we always have $f(\overline{A}) = \overline{f(A)}$.

(ii) Let $a \in \overline{A}$ and $b \in \overline{B}$ and let W be a neighbourhood of a + b. Then there are neighbourhoods W_a, W_b of a and b such that $W_a + W_b \subseteq W$. By the definition of the closure of A (intersection of closed supersets of A) we immediately get that any neighbourhood of a must have a non-empty intersection with A. Thus there exist points $x \in A \cap W_a$ and $y \in B \cap W_b$. Thus

$$x + y \in (A + B) \cap (W_a + W_b) \subseteq (A + B) \cap W_a$$

In particular $(A + B) \cap W \neq \emptyset$. Since W was arbitrary, we get that $a + b \in \overline{A + B}$.

(iii) From the first two points we get

$$\lambda \overline{A} + (1 - \lambda)\overline{A} = \overline{\lambda}\overline{A} + \overline{(1 - \lambda)A} \subseteq \overline{\lambda}\overline{A} + (1 - \lambda)\overline{A} = \overline{A}.$$

We will also need the notion of the convex hull. For K being a subset of a vector space X we define

$$\operatorname{conv}(K) = \left\{ \sum_{i=1}^{n} \lambda_i x_i : x_i \in K, \ \lambda_i \ge 0, \ \sum_{i=1}^{n} \lambda_i = 1, \ n \ge 1 \right\}.$$

This is clearly the smallest convex set which contains K.

Fact 42. Suppose A_1, \ldots, A_n are convex compact subsets of a t.v.s. X. Then $\operatorname{conv}(A_1 \cup \ldots \cup A_n)$ is compact.

Proof. Let

$$S = \{ (s_1, \dots, s_n) : s_i \ge 0, i = 1, \dots, n, s_1 + \dots + s_n = 1 \}.$$

Take $A = A_1 \times \ldots \times A_n$ and define $f : S \times A \to X$ via $f(s, a) = s_1 a_1 + \ldots + s_n a_n$. Take $K = f(S \times A)$. This set is clearly compact as an image of a compact set under continuous map. Clearly $K \subseteq \operatorname{conv}(A_1 \cup \ldots \cup A_n)$. It is also straightforward to check that K is convex. Thus since $A_i \subseteq K$ (take $s_i = 1$) we get $A_1 \cup \ldots \cup A_n \subseteq K$ and by convexity of K we arrive at $\operatorname{conv}(A_1 \cup \ldots \cup A_n) \subseteq K$, which yields $\operatorname{conv}(A_1 \cup \ldots \cup A_n) = K$ and the compactness of $\operatorname{conv}(A_1 \cup \ldots \cup A_n)$ follows. \Box

3.2. Locally convex spaces.

Definition 3. A t.v.s. X is called locally convex (l.c.) if it has a local base whose members are convex.

Theorem 43 (Milman's theorem). Let X be a l.c.t.v.s. and let K be a compact set such that $\overline{\text{conv}}(K)$ is also compact. Then $\exp(\overline{\text{conv}}(K)) \subseteq K$.

Proof. Assume there is $p \in \text{ext}(\overline{\text{conv}}(K))$ such that $p \notin K$. From Fact 38 and Fact 39 we can find a neighbourhood V of 0 such that $(p + \overline{V}) \cap K = \emptyset$. Moreover, by the definition of l.c. spaces we can assume that V is convex. Furthermore, by taking $V \cap (-V)$ instead of V be can assume that V is symmetric (V = -V). By compactness of K there are points $x_1, \ldots, x_n \in K$ such that $K \subseteq \bigcup_{i=1}^n (x_i + V)$. The sets

$$A_i = \overline{\operatorname{conv}}((x_i + V) \cap K) \subseteq \overline{\operatorname{conv}}(K)$$

are closed subsets of compact set and thus they are compact. They are also convex as closures of convex sets (Fact 41 (iii)). Clearly $K \subseteq \bigcup_{i=1}^{n} A_i$. Thus Fact 42, together with the fact that in Hausdorff spaces compact sets are closed, yields

$$\overline{\operatorname{conv}}(K) \subseteq \overline{\operatorname{conv}}\left(\bigcup_{i=1}^{n} A_{i}\right) = \operatorname{conv}\left(\bigcup_{i=1}^{n} A_{i}\right).$$

Since $A_i \in \overline{\operatorname{conv}}(K)$ we also have $\operatorname{conv}(\bigcup_{i=1}^n A_i) \subseteq \overline{\operatorname{conv}}(K)$ and thus

$$\overline{\operatorname{conv}}(K) = \operatorname{conv}\left(\bigcup_{i=1}^{n} A_i\right).$$

In particular, $p = \sum_{i=1}^{n} \lambda_i a_i$, where $a_i \in A_i \subseteq \overline{\text{conv}}(K)$ and $\lambda_i \in [0, 1]$ sum up to 1. By extremality of p in $\overline{\text{conv}}(K)$ we get that there is i such that

$$p \in A_i = \overline{\operatorname{conv}}((x_i + V) \cap K) \subseteq x_i + \overline{V} \subseteq K + \overline{V}.$$

Here the first inclusion follows from the fact that $x_i + \overline{V}$ are closed and convex. Get get that p = k + v, where $k \in K$ and $v \in \overline{V}$. Since \overline{V} is symmetric (since $\overline{V} = \overline{-V} = -\overline{V}$) we get $p + \overline{V} \ni p - v = k \in K$ and thus $(p + \overline{V}) \cap K \neq \emptyset$ contradicting our initial assumption. \Box

Theorem 44 (Bauer's Maximum Principle). Let K be a nonempty compact convex set in some l.c.t.v.s., and let $g: K \to \mathbb{R}$ be a convex upper semi-continuous function. Then g attains its maximum over K at some extreme point of K.

Proof. Step 1. Let $m = \sup_{x \in K} g(x)$. We first prove that $m < \infty$. Indeed, take the sets $M_n\{x \in K : g(x) \ge n\}$. Then M_n are closed (since g is upper semi-continuous) and their intersection is empty. Thus $(M_n^c)_n$ form an open covering of K and so by compactness of K there is a finite sub-cover $(M_{n'}^c)_{n'}$. One of these sets includes the other ones (since the family is decreasing) and thus for some m we have $K = M_m^c$ and thus $M_m = \emptyset$. The assertion follows.

Step 2. Define $M = \{y \in K : g(y) = m\}$. We claim that this set is non-empty. To prove it define closed non-empty sets $M_n = \{x \in K : g(x) \ge m - \frac{1}{n}\}$. We have $M = \bigcap_{n=1}^{\infty} M_n$ and this intersection is non empty (the argument is similar to the above reasoning; assuming empty intersection we get that (M_n^c) is a cover of K which then has an open sub-cover (M_n^c)), but this means that $\bigcap_m (M_m^c)$ is empty, this is not possible since the family is decreasing and all the sets are non-empty).

Step 3. We claim that M is compact. Indeed $M = \{y \in K : g(y) = m\} = \{y \in K : g(y) \ge m\}$, which is closed as g is upper semi-continuous. Since $M \subseteq K$ and K is compact, the assertion follows (recall that closed subsets of compact sets in Hausdorff spaces are compact).

Step 4. We shall show that M is extremal in K, that is, whenever $x \in M$ is written in the form $x = \lambda x_1 + (1 - \lambda)x_2$ for some $x_1, x_2 \in K$, we must have $x_1, x_2 \in M$. Indeed, for such representation we gave, by convexity of g,

$$m = g(x) = g(\lambda x_1 + (1 - \lambda)x_2) \le \lambda g(x_1) + (1 - \lambda)g(x_2) \le \lambda m + (1 - \lambda)m = m$$

Thus $g(x_1) = g(x_2) = m$ and so $x_1, x_2 \in M$.

Step 5. We claim that if $M \subseteq K$ is extremal in K then

(25)
$$\operatorname{ext}(\operatorname{\overline{conv}}(M)) \subseteq M \cap \operatorname{ext}(K).$$

Since M is compact, the inclusion $\operatorname{ext}(\operatorname{\overline{conv}}(M)) \subseteq M$ follows from Milman's theorem (Theorem 43). Note that in order to use this theorem we need to know that $\operatorname{\overline{conv}}(M)$ is compact which is true as it is a close subset of K (this follows from the fact that $M \subseteq K$ and K is closed and convex). Not suppose $x \in \operatorname{ext}(\operatorname{\overline{conv}}(M))$. We shall show that $x \in \operatorname{ext}(K)$. Of course from Milman $x \in M \subseteq K$. Assume that $x = \lambda x_1 + (1 - \lambda) x_2$ for some $x_1, x_2 \in K$. Since M is extremal in K, we get that $x_1, x_2 \in M \subseteq \operatorname{\overline{conv}}(M)$. By the extremality of x in $\overline{\operatorname{conv}}(M)$ we infer that $x_1 = x_2 = x$. This proves that $x \in \operatorname{ext}(K)$.

Step 6. Now it suffices to use Krain-Milman theorem to claim that $ext(\overline{conv}(M))$ is nonempty (together with the fact that the closure of a convex set is convex, which is Fact 41(iii)). Thus from (25) there is an element in $M \cap ext(K)$, that is an extremal point such that g(x) = m.

4. KLS LOCALIZATION PROOF

Let K be a compact set in \mathbb{R}^n . The function $f: K \to \mathbb{R}$ is called upper semi-continuous if for any sequence $(x_n) \subseteq K$ converging to some point $x \in K$ we have

$$\limsup_{n \to \infty} f(x_n) \le f(x)$$

Equivalently, for any y the set $\{f \ge y\}$ is closed. In fact we shall use this second definition to deal with functions defined on arbitrary topological spaces.

Theorem 45. Let $\mathcal{P}(K)$ be the set of regular Borel probability measures supported in the compact set $K \subseteq \mathbb{R}^n$. Suppose $f: K \to \mathbb{R}$ and $\phi: \mathcal{P}(K) \to \mathbb{R}$ are upper semi-continuous. Define $P_f \subseteq \mathcal{P}(K)$ via

$$P_f = \left\{ \mu : \ \mu \text{ is a log concave probability measure in } K \text{ with } \int f d\mu \ge 0 \right\}.$$

Then $\sup_{\mu \in P_f} \phi(\mu)$ is attained on $ext(conv(P_f))$.

Before we give a proof of this fact let us discuss certain preparatory facts. The set $\operatorname{conv}(P_f)$ consists of certain probability measures. We treat it as a subset of the linear space of Borel regular Radon measures, which is known to be the dual of C(K), the space of continuous function on K with the sup norm

$$\|f\| = \sup_{x \in K} f(x).$$

On the dual $C(K)^*$ we can consider a norm given by

$$\|\mu\| = \sup\{\mu(f): f \in C(K), \|f\| \le 1\}.$$

However, we shall equip $C(K)^*$ with the so-called weak-* topology, which is the smallest topology such that for any $f \in C(K)$ the pointwise evaluation functionals $T_f : C(K)^* \to \mathbb{R}$ given by $T_f(\mu) = \mu(f)$ are continuous. Due to the celebrated Banach-Alaoglu theorem for any normed space X the unit ball (in the dual norm) in the dual space X^* is weak-* compact. Thus, the unit ball $B_{C(K)^*}$ in the space $C(K)^*$ is weak-* compact. The unit ball in this case consists of all measures $\mu \in C(K)^*$ satisfying

$$\left| \int f \mathrm{d}\mu \right| \le 1$$
, for all f with $||f|| \le 1$.

In particular, we trivially have $P_f \subseteq B_{C(K)^*}$.

Note that $B_{C(K)^*}$ is closed in weak-* topology. Indeed, the set of measures satisfying $|\int f d\mu| \leq 1$ for fixed f is the same as the set of functionals satisfying $|\mu(f)| \leq 1$ (in the functional analytic notation), which is closed due to the definition of weak-* topology. We now intersect these sets for all f with $||f|| \leq 1$ and get that $B_{C(K)^*}$ is closed. As a consequence, since $B_{C(K)^*}$ is convex and closed, the set $\overline{\operatorname{conv}}(P_f)$ is a closed subset of a compact set $B_{C(K)^*}$ and thus it is itself compact (general easy fact from topology saying that a closed subset of a compact set is compact).

The claim that the set $\mathcal{P}(K)$ is weak-* compact. Of course it is a subset of $B_{C(K)^*}$. The only thing we have to show is that $\mathcal{P}(K)$ is closed. Let A be a closed subset of K. Define $f_{\varepsilon,A}(x) = (1 - \operatorname{dist}(x, A)/\varepsilon)_+$. From the definition of weak-* topology the set of measures satisfying $\int f_{\varepsilon,A} d\mu \in [0, 1]$ is closed. Intersecting this for any $\varepsilon > 0$ and any A shows that the set of measures in $C(K)^*$ satisfying $\int f_{\varepsilon,A} d\mu \in [0, 1]$ for any A and any $\varepsilon > 0$ is closed. We claim that this set is actually equal to the set of measure μ in $C(K)^*$ satisfying $\mu(A) \in [0, 1]$ for any Borel set $A \subseteq K$. Indeed, for closed sets, by the Lebesgue dominated convergence theorem for there measures, we have

$$\lim_{\varepsilon \to 0^+} \int f_{\varepsilon,A} \mathrm{d}\mu = \int \lim_{\varepsilon \to 0^+} f_{\varepsilon,A} \mathrm{d}\mu = \mu(A) \in [0,1].$$

Thus, for all closed (and open) sets A we have $\mu(A) \in [0, 1]$. The assertion for general Borel sets follows form the regularity of μ . If we further intersect our set with the set of measures satisfying $\mu(K) = 1$ (which is again closed as a preimage of the evaluation function for $f \equiv 1$), we get that the set $\mathcal{P}(K)$ is closed in $C(K)^*$.

We would like to show that also P_f is compact. Consider continuous functions $f, g, h : K \to \mathbb{R}$. The function

$$\Psi(\mu) = \int h \mathrm{d}\mu - \left| \int f \mathrm{d}\mu \right|^{\lambda} \left| \int g \mathrm{d}\mu \right|^{1-\lambda}$$

is the superposition of the maps

$$\mu \to \left(\int f \mathrm{d}\mu, \int g \mathrm{d}\mu, \int h \mathrm{d}\mu\right)$$
 and $\Phi(x, y, z) = z - |x|^{\lambda} |y|^{1-\lambda}$

and thus it is continuous. Thus $\Psi^{-1}([0,\infty))$ is closed. If we intersect this set with the closed $\mathcal{P}(K)$, we get that for any continuous non-negative functions f, g, h the set of probability Radon measures satisfying

$$\int h \mathrm{d}\mu \ge \left(\int f \mathrm{d}\mu\right)^{\lambda} \left(\int g \mathrm{d}\mu\right)^{1-\lambda}$$

is closed and therefore compact. Therefore, the Radon probability measures satisfying the assertion of Prékopa-Leindler inequality for for continuous functions, that is

$$h(\lambda x + (1 - \lambda)y) \ge f(x)^{\lambda} f(y)^{1-\lambda} \implies \int h d\mu \ge \left(\int f d\mu\right)^{\lambda} \left(\int g d\mu\right)^{1-\lambda}$$

form a compact set. These are precisely the log-concave measures. Indeed, log-concave measures satisfy the Prékopa-leindler inequality. On the other hand, if Prékopa-Leindler is satisfied for continuous f, g, h and the measure μ is regular, then it is also satisfied for general Borel-measurable functions and thus taking standard function $h = \mathbf{1}_{\lambda A + (1-\lambda)B}, f = \mathbf{1}_A$ and $g = \mathbf{1}_B$ yields the inequality

$$\mu(\lambda A + (1 - \lambda)B) \ge \mu(A)^{\lambda}\mu(B)^{1-\lambda},$$

which gives the log-concavity of μ (due to the celebrated result of Borel, which we shall not discuss here, this is the same as the set P_f used by us). We have shown the compactness of log-concave measures on K. Now, the compactness of P_f follows by observing that the set of measures in $\mathcal{P}(K)$ satisfying $\int f d\mu \geq 0$ for a fixed upper semi-continuous function is closed. This would follow if we could prove the semi-continuity of the functional $\mu \mapsto \int f d\mu$. Upper semi-continuous functions are known to be monotone non-increasing limits of continuous functions. Therefore there are continuous functions $f_n \searrow f$. If μ is a Radon probability measure then $\int f d\mu \geq 0$ is equivalent to $\int f_n d\mu \geq 0$ for any $n \geq 1$, due to the Lebesgue dominated convergence theorem. The set

$$\bigcap_{i=1}^{\infty} \left\{ \mu : \int f_n \mathrm{d}\mu \ge 0 \right\}$$

is compact as the intersection of compact sets and thus we deduce that P_f closed and therefore compact.

We are now ready go give a proof of Theorem 45.

Proof of Theorem 45. We already know that P_f is compact. Applying Bauer's principle (Theorem 44) to the convex compact set $\overline{\operatorname{conv}}(P_f)$ we get that Φ attains its maximum at $\operatorname{ext}(\overline{\operatorname{conv}}(P_f))$. We will show that $\operatorname{ext}(\overline{\operatorname{conv}}(P_f)) \subseteq \operatorname{ext}(\operatorname{conv}(P_f))$. Indeed, suppose $x \in \operatorname{ext}(\overline{\operatorname{conv}}(P_f))$. By Milman's theorem (Theorem 43) we have $x \in P_f$. Suppose that $x = \lambda x_1 + (1 - \lambda) x_2, \lambda \in (0, 1), x_1, x_2 \in \operatorname{conv}(P_f)$. Obviously $x_1, x_2 \in \overline{\operatorname{conv}}(P_f)$ and by extremality of x in $\overline{\operatorname{conv}}(P_f)$ we get $x = x_1 = x_2$. Thus, $x \in \operatorname{ext}(\operatorname{conv}(P_f))$. The description of possible maximizers follows from Theorem 46.

Remark 7. In the above reasoning it is important to know that $x \in P_f$. Otherwise we would not be able to use the definition of extremality, which requires the point to be in the convex set we are dealing with. To understand it better one can attempt to prove a false statement that $A \subseteq B$ (A, B - convex) implies $\text{ext}(B) \subseteq \text{ext}(A)$.

4.1. Description of extreme points. Let us mention that the set of extreme points of the convex hull of log-concave measures on K is the set of all Dirac masses. Indeed, if μ is log-concave and μ is not a Dirac mass, then there is a hyperplane H dividing \mathbb{R}^n into half-spaces H^+ and H^- such that $\mu_{|H^+}$ and $\mu_{|H^-}$ are non-zero measures and $\mu(H) = 0$. Thus

$$\mu = \mu(H^+) \frac{\mu_{|H^+}}{\mu(H^+)} + \mu(H^-) \frac{\mu_{|H^-}}{\mu(H^-)}$$

which is a non-trivial convex combination of log-concave probabilities $\mu_{|H^+}/\mu(H^+)$ and $\mu_{|H^-}/\mu(H^-)$. Thus μ is not extreme.

Our goal is to characterize extreme points of conv P_f . A segment $[a, b] \subset \mathbb{R}^n$ is the set $\{a + t(b - a) : t \in [0, 1]\}$, where $a, b \in \mathbb{R}^n$. We shall discuss the following theorem.

Theorem 46. Let ν be an extreme point of $\operatorname{conv}(P_f)$. Then one of the following holds

- (i) ν is a Dirac masses at point x such that $f(x) \ge 0$,
- (ii) ν is supported on a segment $[a, b] \subset K$ such that on that segment the density of ν is log-affine, $\int f d\nu = 0$ and

(26)
$$\int_{a}^{x} f d\nu > 0$$
, for all $x \in (a, b)$ or $\int_{x}^{c} f d\nu > 0$, for all $x \in (a, b)$.

Proof. Let us assume that ν is an extreme point in $\operatorname{conv}(P_f)$ which is not a Dirac mass. We will prove that it is of the form (ii). Clearly we have $\nu \in P_f$ since otherwise by the definition of convex hull it is a non-trivial combination of elements of P_f . Let G be the least affine subspace containing the support of ν .

Step 1. We will prove that dim G = 1. Suppose that dim $G \ge 2$. Let x_0 be any interior point (in G) of the support of ν (the support of ν has interior point as it is a convex set of full dimension (in G). Let E be a two dimensional subspaces such that $x_0 + E \subseteq G$. Take a unit circle in $S^1(E)$ in E and for any $u \in S^1(E)$ define hyperplane H_u and half-spaces H_u^+ and H_u^- by

$$H_u = \{ x \in G : \langle x - x_0, u \rangle = 0 \}, \qquad H_u^{\pm} = \{ x \in G : \langle x - x_0, \pm u \rangle \ge 0 \}.$$

Define $\varphi: S^1(E) \to \mathbb{R}$ by $\varphi(u) = \int_{H_u^+} f d\nu - \frac{1}{2} \int f d\nu$. Clearly $\varphi(u) + \varphi(-u) = 0$, which follows from the fact that $H_{-u}^+ = H_u^-$. Thus $\varphi(-u) = -\varphi(u)$ which means that $\varphi(u)$ and $\varphi(-u)$ are either both zero or have opposite signs. The usual Darboux principle together with the continuity of φ (which follows from the fact that $\nu(H_u) = 0$) shows that there is $\varphi(u_0) = 0$. The choice of x_0 ensures that $\nu(H_{u_0}^+) > 0$ and $\nu(H_{u_0}^-) > 0$. Since $\varphi(u_0) = \varphi(-u_0) = 0$ we get that $\varphi(u) = \int_{H_u^{\pm}} f d\nu = \frac{1}{2} \int f d\nu$. Thus, the measures

$$\mu_{|H_{u_0}^+}/\mu(H_{u_0}^+), \qquad \mu_{|H_{u_0}^-}/\mu(H_{u_0}^-)$$

are probability measure belonging to P_f and

$$\mu = \mu(H_{u_0}^+) \frac{\mu_{|H_{u_0}^+}}{\mu(H_{u_0}^+)} + \mu(H_{u_0}^-) \frac{\mu_{|H_{u_0}^-}}{\mu(H_{u_0}^-)},$$

which contradicts extremality of ν .

Step 2. We can therefore assume, without loss of generality, that ν is supported on $[a, b] \subset \mathbb{R}$. If the continuous function $x \mapsto \int_a^x f d\nu$ has a sign on (a, b) then it is either positive or negative. In the former case the first condition in (26) holds true, whereas in latter case the second condition has to be satisfied as

$$\int_{x}^{c} f d\nu = \int f d\nu - \int_{a}^{x} f d\nu \ge - \int_{a}^{x} f d\nu \ge 0.$$

Thus, suppose that for some $x \in (a, b)$ we have $\int_a^x f d\nu = 0$. Since then $\int_x^c f d\nu \ge 0$ we get that the measures $\nu_1 = \nu_{|[a,x]}/\nu([a,x])$ and $\nu_2 = \nu_{|[x,c]}/\nu([x,c])$ belong to P_f and satisfy $\nu = \nu[a,x]\nu_1 + \nu[x,c]\nu_2$, which again leads to contradiction. To prove that $\int f d\nu = 0$ let us assume that $\int f d\nu > 0$. By Darboux principle there is $x \in (a,b)$ such that $\int_a^x f d\nu = \frac{1}{2} \int f d\nu > 0$. The also $\int_x^c f d\nu = \frac{1}{2} \int f d\nu > 0$ and defining ν_1 and ν_2 as before again leads to contradiction.

Step 3. We shall finally prove that ν is log-affine. Without loss of generality we can assume that $\int_a^x f d\nu > 0$ for all $x \in (a, b)$. Let ψ be the density of ν . Take any $c \in (a, b)$ and define $\varphi_{\alpha}(x) = \frac{1}{2}\psi(c)e^{\alpha(x-c)}$. Consider the measures

$$d\mu_{\alpha} = (\psi - \varphi_{\alpha})_{+} dx, \qquad d\nu_{\alpha} = \min\{\psi, \varphi_{\alpha}\} dx.$$

Note that since $\varphi_{\alpha}(x) = \frac{1}{2}\psi(c) < \psi(c)$, the measure μ_{α} is non-zero. It is clear that ν_{α} is log-concave, as the maximum of convex functions is convex. We claim that also μ_{α} is log-concave. To check it we observe that the support of μ_{α} is an interval (the inequality $\psi \geq \varphi_{\alpha}$ is equivalent to $\ln \psi - \ln \varphi_{\alpha} \geq 0$, where the left hand side is concave). Thus its support the measure has density $\psi - \varphi_{\alpha} = \varphi_{\alpha}(e^{-V} - 1)$, where $V = -\ln(\psi/\varphi_{\alpha})$ is convex and non-positive. We check that $g = e^{-V} - 1$ is log-concave. We are to check that $gg'' \leq (g')^2$. This is equivalent to

$$(e^{-V} - 1)((V')^2 - V'')e^{-V} \le (V')^2 e^{-2V},$$

which is the same as $V''(1 - e^{-V}) \leq (V')^2$. This is true as the left hand side is non-positive. Thus μ_{α} is log-concave.

Since $\int f d\nu = 0$ and $\int_a^c f d\nu > 0$, we have (by using Lebesgue dominated convergence theorem)

$$\lim_{\alpha \to -\infty} \int f d\nu_{\alpha} = \int_{a}^{c} f d\nu > 0, \qquad \lim_{\alpha \to +\infty} \int f d\nu_{\alpha} = \int_{c}^{b} f d\nu < 0.$$

Thus, by continuity of $\alpha \to \int f d\nu_{\alpha}$ there is α_0 such that $f d\nu_{\alpha_0=0}$. Clearly $\mu_{\alpha} + \nu_{\alpha} = \nu$ and thus $\int f d\mu_{\alpha_0} = \int f d\nu = 0$. Take

$$\nu_1 = \frac{\mu_{\alpha_0}}{1-\lambda}, \quad \nu_2 = \frac{\nu_{\alpha_0}}{\lambda}, \quad \text{where} \quad \lambda = \nu_{\alpha_0}[a,b], \quad 1-\lambda = \nu[a,b] - \nu_{\alpha_0}[a,b] = \mu_{\alpha_0}[a,b].$$

We get $\nu = \lambda \nu_1 + (1 - \lambda)\nu_2$, where ν_1 and ν_2 are probability measures in P_f . This is a contradiction.

We are ready to deduce the so-called four function theorem.

Theorem 47. Let $f_1, f_2, f_3, f_4 : \mathbb{R}^n \to \mathbb{R}$ be nonnegative and such that f_1, f_2 are upper semi-continuous and f_3, f_4 are lower semi-continuous. Suppose $\alpha, \beta > 0$. Then the inequality

$$\left(\int f_1 \mathrm{d}\mu\right)^{\alpha} \left(\int f_2 \mathrm{d}\mu\right)^{\beta} \le \left(\int f_3 \mathrm{d}\mu\right)^{\alpha} \left(\int f_4 \mathrm{d}\mu\right)^{\beta}$$

holds true for any log-concave measure μ if and only if it holds for Dirac masses and log-affine measures supported on one-dimensional segments.

Proof. By considering $\mu_k = \mu_{|kB_2^n}$ one can assume that μ is compactly supported on some convex compact set K. By considering $f_3 + \frac{1}{n}$ instead of f_3 we can assume that $f_3 > 0$. Define

$$f = f_1 - \frac{\int f_1 d\mu}{\int f_3 d\mu} f_3, \qquad \Phi(\theta) = \left(\frac{\int f_1 d\mu}{\int f_3 d\mu}\right)^{\frac{1}{\beta}} \int f_2 d\theta - \int f_4 d\theta.$$

The functional Φ is affine and upper semi-continuous. Indeed by upper semi-continuity of f_2 and lower semi-continuity of f_4 we have

$$\limsup_{n \to \infty} \int f_2 d\mu_n \le \int f_2 d\mu, \qquad \liminf_{n \to \infty} \int f_4 d\mu_n \ge \int f_4 d\mu$$

whenever $\mu_n \Rightarrow \mu$. Clearly $\mu \in P_f$. By Theorem ?? we get that there is $\nu \in P_f$ of the special form described in Theorem 46 such that $\Phi(\mu) \leq \Phi(\nu)$. Thus,

$$\left(\frac{\int f_1 d\mu}{\int f_3 d\mu}\right)^{\frac{\alpha}{\beta}} \int f_2 d\mu - \int f_4 d\mu \le \left(\frac{\int f_1 d\mu}{\int f_3 d\mu}\right)^{\frac{\alpha}{\beta}} \int f_2 d\nu - \int f_4 d\nu$$
$$\le \left(\frac{\int f_1 d\nu}{\int f_3 d\nu}\right)^{\frac{\alpha}{\beta}} \int f_2 d\nu - \int f_4 d\nu \le 0,$$

where the second inequality follows form $\int f d\nu \ge 0$ and the last from the assumptions of the theorem. The assertion follows.

We shall give several examples of the use of localization lemma.

Example 4. Let us prove the following theorem.

Theorem 48. Suppose Hess $V \ge c^2 I$ on \mathbb{R}^n , where c > 0. Let μ be a probability measure with density e^{-V} . Then for all measurable sets A we have

(27)
$$\mu(A_h) \ge \Phi\left(\Phi^{-1}(\mu(A)) + ch\right),$$

where $\Phi(s) = (2\pi)^{-1/2} \int_{-\infty}^{s} e^{-x^2/2} dx$ and $A_h = \{x : d(x, A) < h\}.$

Proof. Step 1. It is enough to consider c = 1. Indeed, let us take the measure $\nu(A) = \mu(A/c)$. Then ν has density $e^{-V(y/c)}c^{-n}$ and thus it satisfies our assumption with c = 1. Since $(\frac{1}{c}A)_h = \frac{1}{c}A_{ch}$ we get, by taking $\frac{1}{c}A$ instead of A that

$$\Phi\left(\Phi^{-1}(\nu(A)) + ch\right) \le \mu((\frac{1}{c}A)_h) = \mu(\frac{1}{c}A_{ch}) = \nu(A_{ch}).$$

Taking h instead of ch finishes the argument.

Step 2. To deal with the case c = 1 take a compact convex set A and a number $m \in (0, 1)$. It suffices to show that $\mu(A) \ge m$ implies $\mu(A_h) \ge \Phi(\Phi^{-1}(m) + h)$. Consider

$$f = \varphi(\mathbf{1}_A - m), \qquad g = -\varphi(\mathbf{1}_{A_h} - \Phi(\Phi^{-1}(m) + h)).$$

Here $A_h = \{x : d(x, A) < h\}$ and thus A_h is open. Thus both f and g are upper semicontinuous. The inequality $\mu(A) \ge m$ is equivalent to $\int f d\mu \ge 0$, where μ is log-concave. The functional $\Phi(\mu) = \int g d\mu$ is upper semi-continuous. Our goal is to show that $\Phi(\mu) \le 0$ on the set P_f . By Theorem ?? the maximal value of this functional is attained on some extreme point of conv (P_f) . Therefore it suffices to show the inequality $\Phi(\nu) \le 0$ for $\nu \in \text{ext}(\text{conv}(P_f))$, that is for ν being log-affine on segments $[a, b] \subset \mathbb{R}^n$ (and for Dirac masses in which case the inequality is obvious as $\delta_x \in P_f$ implies $x \in A$ and thus $x \in A_h$ which gives $\Phi(\delta_x) \le 0$).

Step 3. Let l be the line containing [a, b]. Since $(A \cap l)_h \subseteq A_h \cap l$, we can assume that A is a subset of the real line and reduce the problem to the case n = 1. We show that every probability measure μ on the real line whose density is of the form $f = \rho \varphi$ is a contraction of γ_1 . The non-decreasing map T transporting γ_1 onto μ satisfies $\mu(-\infty, T(x)) = \Phi(x)$, that is $T(x) = F^{-1}(\Phi(x))$. Computing the derivative gives

$$T'(x) = \frac{\Phi'(x)}{F'(F^{-1}(\Phi(x)))} = \frac{\varphi(x)}{f(F^{-1}(\Phi(x)))}$$

We would like to show that $T'(x) \leq 1$, which is equivalent to $\varphi(x) \leq f(F^{-1}(\Phi(x)))$. Taking $p = \Phi(x)$ we can rewrite it in the form $\varphi(\Phi^{-1}(p)) \leq f(F^{-1}(p))$. We shall prove something more general, namely that every finite measure (non-necessarily probability measure) satisfies the implication

$$\mu(-\infty, x_0) \ge p, \quad \mu(x_0, \infty) \ge 1 - p, \quad p \in (0, 1) \qquad \Longrightarrow \qquad f(x_0) \ge \varphi(\Phi^{-1}(p)).$$

To this end we first assume that our assertion is true for ρ being log-affine on \mathbb{R} . We shall prove it for general ρ . Indeed, let l(x) be the tangent line to the graph of a convex function $U = -\ln \rho$ at x_0 . Define $\rho_0 = e^{-l} \ge \rho$ and let μ_0 be the measure with density $\rho_0 \varphi$. If μ, p and x_0 satisfy $\mu(-\infty, x_0) \ge p$, $\mu(x_0, \infty) \ge 1 - p$ for some $p \in (0, 1)$, then also μ_0 satisfies the same condition as $\rho_0 \ge \rho$. Thus $f(x_0) = f_0(x_0) \ge \varphi(\Phi^{-1}(p))$.

So, it is enough to assume that $f(x) = Ce^{\lambda x}\varphi(x)$. The conditions

$$\mu(-\infty, x_0) = \int_{-\infty}^{x_0} C e^{\lambda s} \varphi(s) \mathrm{d}s = C e^{\lambda^2/2} \Phi(x_0 - \lambda) \ge p,$$

$$\mu(x_0, \infty) = \int_{x_0}^{\infty} C e^{\lambda s} \varphi(s) \mathrm{d}s = C e^{\lambda^2/2} (1 - \Phi(x_0 - \lambda)) \ge 1 - p.$$

are equivalent to

$$C \ge e^{-\lambda^2/2} \max\left\{\frac{p}{\Phi(x_0 - \lambda)}, \frac{1 - p}{1 - \Phi(x_0 - \lambda)}\right\}$$

Thus

$$f(x_0) = Ce^{\lambda x_0}\varphi(x_0) \ge e^{-\lambda^2/2}e^{\lambda x_0}\varphi(x_0) \max\left\{\frac{p}{\Phi(x_0-\lambda)}, \frac{1-p}{1-\Phi(x_0-\lambda)}\right\}$$
$$= \varphi(x_0-\lambda) \max\left\{\frac{p}{\Phi(x_0-\lambda)}, \frac{1-p}{1-\Phi(x_0-\lambda)}\right\}.$$

It is therefore enough to show that for every $y \in \mathbb{R}$ we have

$$\varphi(y) \max\left\{\frac{p}{\Phi(y)}, \frac{1-p}{1-\Phi(y)}\right\} \ge \varphi(\Phi^{-1}(p)).$$

For $y \leq \Phi^{-1}(p)$ the inequality reduces to $p\varphi(y)/\Phi(y) \geq \varphi(\Phi^{-1}(p))$. Here $y = \Phi^{-1}(p)$ give equality. Thus, it suffices to observe that the function $\varphi(y)/\Phi(y) = (\log \Phi(y))'$ is decreasing. This follows form the fact that $\Phi(y)$ is concave as the tail of log-concave measure (see Lemma ??). The case $y \geq \Phi^{-1}(p)$ follows by the same argument.

Step 4. Now we show that the validity of (27) is reserved under contractions. In general, we shall show that every inequality of the form

$$\mu(A_h) \ge \Psi(\mu(A), h)$$

is preserved. Indeed, suppose μ is an image of μ_0 under a map T with Lipschitz norm at most 1. Then

$$\mu(A_t) = \mu(A + tB_2^n) = \mu_0(T^{-1}(A + tB_2^n)) \ge \mu_0(T^{-1}(A) + tB_2^n)$$

$$\ge \Psi(\mu_0(T^{-1}(A)), h) = \Psi(\mu(T^{-1}(A)), h).$$

Here the first inequality follows from the inclusion $T^{-1}(A) + tB_2^n \subseteq T^{-1}(A + tB_2^n)$. To show it observe that it is equivalent to $T(T^{-1}(A) + tB_2^n) \subseteq A + tB_2^n$. Now take a point $x \in T^{-1}(A)$. We have to show that $T(x + ty) \in A + tB_2^n$ for every $y \in B_2^n$. In other words, we shall show that d(T(x + ty), A) < t. This is true as $T(x) \in A$ and $d(T(x + ty), T(x)) \leq t|y| < t$.

Step 5. The theorem is now established as γ_1 satisfies (27) with c = 1 (Gaussian isoperimetric inequality). If one wants to get (27) only for convex sets, it is enough to check it for γ_1 and interval on the real line. Let us do this. Let us fix $m \in (0, 1)$ and consider intervals [a, b] such that

$$\gamma_1[a,b] = (2\pi)^{-1/2} \int_a^b e^{-x^2/2} \mathrm{d}x = m.$$

We would like to maximize

$$\psi(a) := \gamma_1([a,b]_h) = \gamma_1[a-h,b+h] = (2\pi)^{-1/2} \int_{a-h}^{b+h} e^{-x^2/2} \mathrm{d}x.$$

Let us consider b = b(a) as a function of a. By symmetry we can assume that $a \leq -b(a)$. Differentiating the constrain gives $e^{-a^2/2} = e^{-b(a)^2/2}b'(a)$. We have

$$\psi'(a) = e^{-(b(a)+h)^2/2}b'(a) - e^{-(a-h)^2/2} = e^{-(b(a)+h)^2/2}e^{-a^2/2}e^{b(a)^2/2} - e^{-(a-h)^2/2}$$
$$= e^{-a^2/2}e^{-h^2/2}(e^{-bh} - e^{ah}) \ge 0.$$

Thus, the minimum is achieved for $a = -\infty$ (which corresponds to half-line) and the maximum for a = -b(a), which give a symmetric interval.

Example 5. Using localization one can reduce proving Brunn-Minkowski inequality in \mathbb{R}^n to the case n = 1. Indeed, suppose we want to show that a probability measure μ with log-concave density supported on some affine subspace of \mathbb{R}^n satisfies

$$\mu(\lambda A + (1 - \lambda)B) \ge \mu(A)^{\lambda}(\mu(B))^{1 - \lambda}, \qquad \lambda \in [0, 1]$$

for every compact sets A, B in \mathbb{R}^n . Let us use Theorem 47 with $f_1 = \mathbf{1}_A$, $f_2 = \mathbf{1}_B$ and $f_3 = f_4 = \mathbf{1}_{(\lambda A + (1-\lambda)B)\varepsilon}$. Here the ε -enlargement is considered to be the open enlargement and thus $f_3 = f_4$ is lower semi-continuous. Taking the limit $\varepsilon \to 0^+$ recovers the desired

inequality due to the continuity of measure. Now, Theorem 47 allows us to reduce our inequality (with $\varepsilon > 0$) to the case of μ being log-affine on some segments $[a, b] \subseteq$.

In this case let us prove the inequality for A, B being convex. Let l be the line containing [a, b]. Since $\lambda(A \cap l) + (1 - \lambda)(B \cap l) \subseteq (\lambda A + (1 - \lambda)B) \cap l$, we can assume that A, B are intervals on the real line, which reduces the problem to the case n = 1. In this case we will prove the desired inequality already for $\varepsilon = 0$. By scaling and translating we can assume that a = 0 and b = 1. The inequality is invariant under multiplying μ by a constant, so the assumption of μ being a probability measure is not needed. Therefore, it is enough to consider $d\mu(x) = e^{\alpha x} \mathbf{1}_{[0,1]}(x), \alpha \in \mathbb{R}$. Note that one can assume that $\inf(A \cup B) = 0$ and $\sup(A \cup B) = 1$, since otherwise we can truncate the support of the measure μ without changing the inequality. We therefore have to consider only two cases: A = [0, c], B = [d, 1] and A = [0, 1], B = [c, d].

In the first case we have $\lambda A + (1 - \lambda)B = [(1 - \lambda)d, 1 - \lambda + \lambda c]$. Thus, we are to show

$$\int_{(1-\lambda)d}^{1-\lambda+\lambda c} e^{\alpha x} \mathrm{d}x \ge \left(\int_0^c e^{\alpha x} \mathrm{d}x\right)^\lambda \left(\int_d^1 e^{\alpha x} \mathrm{d}x\right)^{1-\lambda}$$

This is equivalent to

$$\left|e^{\alpha(1-\lambda+\lambda c)} - e^{\alpha(1-\lambda)d}\right| \ge \left|e^{\alpha c} - 1\right|^{\lambda} \left|e^{\alpha} - e^{\alpha d}\right|^{1-\lambda}.$$

Dividing this by $e^{\alpha(1-\lambda)d}$ gives

$$|e^{\alpha(1-\lambda+\lambda c-(1-\lambda)d)}-1| \ge |e^{\alpha c}-1|^{\lambda} |e^{\alpha(1-d)}-1|^{1-\lambda}.$$

Taking $x = e^{\alpha c}$ and $y = e^{\alpha(1-d)}$ gives $|x^{\lambda}y^{1-\lambda} - 1| \ge |x - 1|^{\lambda}|y - 1|^{1-\lambda}$, where $x, y \ge 1$ (if $\alpha \ge 0$) or $x, y \le 1$ (if $\alpha \le 0$). Since this inequality is invariant under changing $x \to 1/x$ and $y \to 1/y$, we can assume that $x, y \ge 1$, in which case the inequality is simply $\psi(\lambda a + (1-\lambda)b) \ge \lambda \psi(a) + (1-\lambda)\psi(b)$, where $\psi(x) = \ln(e^x - 1)$ and $a = \ln x, b = \ln y$. The concavity of ψ can we checked by observing that $\psi'(x) = 1/(1 - e^{-x})$, which is decreasing.

The second case A = [0, 1], B = [c, d] leads to the same computations.

Example 6. We shall prove the following theorem.

Theorem 49. For every symmetric convex compact set K and a log-concave measure μ on \mathbb{R}^n we have

$$\mu(tK) \ge 1 - (1 - \mu(K))^{\frac{t+1}{2}}, \qquad t \ge 1.$$

Clearly the above inequality cannot be true for arbitrary non-symmetric convex set K. This creates some difficulties in proving the inequality. We need to modify the definition of the dilation tK. Let us set

$$K^t = \{x \in \mathbb{R}^n : \text{ there is an interval } I \ni x \text{ such that } |I| \le \frac{t+1}{2} |K \cap I|\}.$$

It turns out that with this definition one can prove the above inequality for general Borel sets K. Here, for simplicity, we shall give a proof for convex K. We will need the following lemma.

Lemma 14. Let K be a convex compact set. Then for every $t \ge 1$ we have

$$K^{t} = K + \frac{t-1}{2}(K-K) = \frac{t+1}{2}K + \frac{t-1}{2}(-K).$$

In particular, if K is symmetric then $K^t = tK$.

Proof. The second inequality follows by convexity. To prove the first one we prove two inclusions. Suppose $x \in K^t \setminus K$. Then there is a segment [b, x] such that $[b, x] \cap K = [b, c]$ and $|x - b| \leq \frac{t+1}{2}|b - c|$. Since c lies between b and x, we can write $c = (1 - \lambda)b + \lambda x$ for some $\lambda \in (0, 1)$. Plugging this into this inequality give $1/\lambda \leq (t + 1)/2$. Thus,

$$x = c + \left(\frac{1}{\lambda} - 1\right)(c - b) \in K + \left(\frac{1}{\lambda} - 1\right)(K - K) \subseteq K + \frac{t + 1}{2}(K - K).$$

To prove the reverse inclusion let us assume that $x \in \frac{t+1}{2}K + \frac{t-1}{2}(-K)$. We can assume that $x \notin K$. Let us write $x = \frac{t+1}{2}c + \frac{t-1}{2}(-b)$ where $c, b \in K$. Let d be such that $[b, x] \cap K = [b, d]$. Since $x = b + \frac{t+1}{2}(c-b)$, the point x is on the line joining b and c, and $|b-c| \leq |b-x|$. Thus,

$$|[b,x]| = |b-x| = \frac{t+1}{2}|b-c| = \frac{t+1}{2}|b-d| = \frac{t+1}{2}|K \cap [b,x]|.$$

This gives $x \in K^t$.

We are ready to give a proof of the inequality

$$\mu(K^t) \ge 1 - (1 - \mu(K))^{\frac{t+1}{2}}, \qquad t \ge 1.$$

for convex sets K. In fact it is enough to prove the inequality $\mu(U) \geq 1 - (1 - \mu(K))^{\frac{t+1}{2}}$ for any open set U containing K^t . Let us take $f = \mathbf{1}_{U^c} - m$ and $\Phi(\mu) = \mu(A)$. Both f and Φ are upper semi-continuous, so they satisfy the assumptions of Theorem ??. Note that $\mu \in P_f$ satisfy $1 - \mu(U) \geq m$. Our goal is to prove that under this constraint we have $(1 - \Phi(\mu))^{\frac{t+1}{2}} \geq m$. Due to Theorem ?? it is enough to check it for log-affine measures ν on segments $[a, b] \subset \mathbb{R}^n$ (note that for Dirac masses this inequality is obvious). This is equivalent to the validity of the inequality

$$\nu(U) \ge 1 - (1 - \nu(K))^{\frac{t+1}{2}}, \quad t > 1.$$

We shall also use the additional information given by Theorem ??, namely that $\int f d\nu = 0$ and either $\int_{[a,x]} f d\nu > 0$ on (a,b) or $\int_{[x,b]} f d\nu > 0$ on (a,b). Without loss of generality we shall assume that the second case holds true. Namely

(28)
$$\nu(U^c \cap [x,b]) > \nu(U^c)\nu([x,b])$$

Clearly we can assume that K is one-dimensional (since $(K \cap l)_t \subseteq K_t \cap l$) and further that $K \subseteq [a, b]$ (since the parts of K outside [a, b] do not contribute to $\mu(K)$ and can only increase $\mu(K^t)$).

We claim that without loss of generality one can assume that $a \in F$. Indeed, let $a' = \inf K$ and suppose that $\nu' = \nu|_{[a',b]}/\nu[a',b]$. We have

$$\nu'(K^c) = \frac{\nu(K^c \cap [a', b])}{\nu[a', b]} = \frac{\nu(K^c) - \nu[a, a']}{1 - \nu[a, a']} \le \nu(K^c).$$

Moreover, form (28) we get

$$\nu'(U^c) = \frac{\nu(U^c \cap [a', b])}{\nu[a', b]} \ge \nu(U^c).$$

Thus, it is harder to verify our assertion for ν' .

Let us then assume (using translation invariance)that a = 0 and $K = [0, c] \subseteq [0, b]$. We can finally get read of U and simply prove

$$\nu(K^t) \ge 1 - (1 - \nu(K))^{\frac{t+1}{2}}, \quad t > 1.$$

Or equivalently,

$$\nu((K^t)^c) \le \nu(K^c)^{\frac{t+1}{2}}, \qquad t > 1.$$

Note that $K^t = \left[-\frac{t-1}{2}c, \frac{t+1}{2}c\right]$. Let $T = \frac{t+1}{2} \ge 1$. We are to show $\nu[Tc, \infty) \le \nu[c, \infty)^T$. Since $[c, \infty) = T^{-1}[Tc, \infty) + (1 - T^{-1})[0, \infty)$ we get by log-concavity of ν

$$\nu[c,\infty) \ge \nu[Tc,\infty)^{T^{-1}} \nu[0,\infty)^{1-T^{-1}} = \nu[Tc,\infty)^{T^{-1}}$$

Example 7. We shall prove the following theorem.

Theorem 50. Suppose K_1, K_2 are two compact disjoint subsets of a compact convex set K in \mathbb{R}^n , such that $d(K_1, K_2) := \inf_{a \in K_1, b \in K_2} d(a, b) > 0$. Then

$$\operatorname{vol}_n(K_1)\operatorname{vol}_n(K_2) \le \frac{M_1(K)}{d(K_1, K_2)\ln 2}\operatorname{vol}_n(K)\operatorname{vol}_n(K \setminus (K_1 \cup K_2)),$$

where

$$M_1(K) = \frac{1}{\operatorname{vol}_n(K)} \int_K |x - b(K)| \mathrm{d}x, \qquad b(K) = \frac{1}{\operatorname{vol}_n(K)} \int_K x \mathrm{d}x$$

In particular, taking μ_K to be the uniform measure on K and $K_1 = A$, $K_2 = (A_{\varepsilon})^c$ so that $d(K_1, K_2) = \varepsilon$ one gets the isoperimetric-type inequality

$$\mu_K(A)(1-\mu_K(A)) \le \frac{M_1(K)}{\ln 2}\mu_K(\partial A).$$

Proof. Let $d(K_1, K_2) = \varepsilon > 0$. Take $f_i = \mathbf{1}_{K_i}$ for i = 1, 2 and $\mathbf{1}_{(K_1 \cup K_2)^c}$. Also, let us take $f_4(x) = |x-u|/\varepsilon \ln 2$, where u is arbitrary vector in \mathbb{R}^n . Then f_1, f_2 are upper semi-continuous and f_3, f_4 are lower semi-continuous. The assertion of the theorem is equivalent to

$$\int_{K} f_1 \mathrm{d}\mu \int_{K} f_2 \mathrm{d}\mu \le \int_{K} f_3 \mathrm{d}\mu \int_{K} f_4 \mathrm{d}\mu$$

for μ being the Lebesgue measure restricted to K. We shall prove the above for every log-concave measure supported on K. Due to Theorem 47 it is enough to consider only the case of μ being log-affine on a segment $[a,b] \subset \mathbb{R}^n$. Clearly, we can assume that [a,b]intersects both K_1 and K_2 . Also, we can assume that u belongs to the line l containing [a,b], since otherwise we can shift the whole picture not changing the integrals of f_1, f_2, f_3 and decreasing the integral of f_4 (|x - u| changes to the length of the distance between xand the orthogonal projection of u onto the line containing the segment [a,b]). We can in fact assume that $u \in [a,b]$, since otherwise we can shift the whole picture along l and again decrease the contribution coming from f_4 . Now we can restrict our attention to the case n = 1 by considering $\tilde{K}_i = K_i \cap l$, i = 1, 2 (note that $d(K_1, K_2) \leq d(\tilde{K}_1, \tilde{K}_2)$). By rescaling and canceling multiplicative constants we can assume that the log-affine density of μ is equal to e^t . Thus, given two disjoint compact subsets K_1, K_2 of [a, b], such that $d(K_1, K_2) = \varepsilon$ we are to show that

$$\int_{K_1} e^t \mathrm{d}t \int_{K_2} e^t \mathrm{d}t \le \frac{1}{\varepsilon \ln 2} \int_{[a,b] \setminus (K_1 \cup K_2)} e^t \mathrm{d}t \int_a^b e^t |t-u| \mathrm{d}t.$$

Let us first assume that K_1, K_2 and $K_3 := [a, b] \setminus (K_1 \cup K_2)$ are intervals. Note that we can assume that $K_3 = [c, d]$ is an interval of length ε (by considering the worst possible case) and is between K_1 and K_2 . Without loss of generality we can assume that $K_1 = [a, c] K_3 = [c, d]$ and $K_2 = [d, b]$ with a < c < d < b and $d - c = \varepsilon$. We are to show

$$\int_{a}^{c} e^{t} \mathrm{d}t \int_{c+\varepsilon}^{b} e^{t} \mathrm{d}t \leq \frac{1}{\varepsilon \ln 2} \int_{c}^{c+\varepsilon} e^{t} \mathrm{d}t \int_{a}^{b} e^{t} |t-u| \mathrm{d}t.$$

Changing variables in the second and third integral we get an equivalent form

$$\int_{a}^{c} e^{t} \mathrm{d}t \int_{\varepsilon}^{b-c} e^{t} \mathrm{d}t \leq \frac{1}{\varepsilon \ln 2} \int_{0}^{\varepsilon} e^{t} \mathrm{d}t \int_{a}^{b} e^{t} |t-u| \mathrm{d}t.$$

The left hand side is equal to

 $(e^{c} - e^{a})(e^{b-c} - e^{\varepsilon}) = e^{b} - e^{\varepsilon}e^{c} - e^{a+b-c} + e^{a+\varepsilon} \le e^{b} - 2e^{\varepsilon/2}e^{(a+b)/2} + e^{a+\varepsilon} = (e^{b/2} - e^{(a+\varepsilon)/2})^{2}.$ by the AM-CM inequality. Clearly without loss of generality we can assume that $a \le u \le b$.

by the AM-GM inequality. Clearly without loss of generality we can assume that $a \le u \le b$ since otherwise one can decrease the right hand side by changing u. We have

$$\int_{a}^{b} e^{t} |t - u| dt = \int_{a}^{u} e^{t} (u - t) dt + \int_{u}^{b} e^{t} (t - u) dt = 2e^{u} - u(e^{a} + e^{b}) + e^{a} (a - 1) + e^{b} (b - 1).$$

The maximum of this function is attained for $u = \ln((e^a + e^b)/2)$ and is equal to

$$ae^a + be^b - (e^a + e^b) \ln\left(\frac{e^a + e^b}{2}\right)$$

Our goal is to verify

$$(e^{b/2} - e^{(a+\varepsilon)/2})^2 \le \frac{e^{\varepsilon} - 1}{\varepsilon \ln 2} \left[ae^a + be^b - (e^a + e^b) \ln\left(\frac{e^a + e^b}{2}\right) \right].$$

Observe that

$$ae^{a} + be^{b} - (e^{a} + e^{b}) \ln\left(\frac{e^{a} + e^{b}}{2}\right) = ae^{a} + be^{b} - (e^{a} + e^{b})(a + \ln\left(\frac{1 + e^{b-a}}{2}\right))$$
$$= (b - a)e^{b} - (e^{a} + e^{b}) \ln\left(\frac{1 + e^{b-a}}{2}\right).$$

Thus, dividing the above inequality by e^a and denoting $z = e^{(b-a)/2}$ we get

$$(z - e^{\varepsilon/2})^2 \le \frac{e^{\varepsilon} - 1}{\varepsilon \ln 2} \left(z^2 \ln z^2 - (1 + z^2) \ln \left(\frac{1 + z^2}{2} \right) \right), \qquad z \ge 1.$$

Clearly the worst case is $\varepsilon = 0$, which leads to

$$(z-1)^2 \ln 2 \le z^2 \ln z^2 - (1+z^2) \ln \left(\frac{1+z^2}{2}\right), \qquad z \ge 1.$$

This is equivalent to

$$-2z\ln 2 \le z^2\ln z^2 - (1+z^2)\ln(1+z^2), \qquad z \ge 1.$$

For z = 1 this is equality. Differentiating and canceling the constant 2 shows that it suffices to prove

$$-\ln 2 \le z \ln z^2 - z \ln(1+z^2), \qquad z \ge 1.$$

Again we have equality for z = 1. We show that the right hand side is monotone in z. Indeed, the derivative is equal to

$$-\ln\left(1+\frac{1}{z^2}\right) + \frac{2}{1+z^2} \ge -\frac{1}{z^2} + \frac{2}{1+z^2} = \frac{z^2-1}{z^2+1} \ge 0, \qquad z \ge 1.$$

To deal with the general case we can assume that $K_3 = [a, b] \setminus (K_1 \cup K_2)$ is open in [a, b]. Thus, it is a union of open intervals. We can assume that these intervals have length at least ε since otherwise both endpoints either belong to K_1 or to K_2 and thus we could add this interval to either K_1 or K_2 , making the inequality tighter. So, let us assume that $K_3 = \bigcup_{i=1}^{k} [c_i, d_i]$, where $|c_i - d_i| \ge \varepsilon$. Using the previous case we get

$$\int_{a}^{c_i} e^t \mathrm{d}t \int_{d_i}^{b} e^t \mathrm{d}t \le \frac{1}{\varepsilon \ln 2} \int_{c_i}^{d_i} e^t \mathrm{d}t \int_{a}^{b} e^t |t - u| \mathrm{d}t, \qquad i = 1, \dots, k.$$

Summing over i we get

$$\sum_{i=1}^{k} \int_{a}^{c_{i}} e^{t} dt \int_{d_{i}}^{b} e^{t} dt \leq \sum_{i=1}^{k} \frac{1}{\varepsilon \ln 2} \int_{c_{i}}^{d_{i}} e^{t} dt \int_{a}^{b} e^{t} |t-u| dt = \int_{K_{3}} e^{t} dt.$$

Now the inequality

$$\sum_{i=1}^{k} \int_{a}^{c_i} e^t \mathrm{d}t \int_{d_i}^{b} e^t \mathrm{d}t \ge \int_{K_1} e^t \mathrm{d}t \int_{K_2} \mathrm{d}t$$

follows from the fact that every point $x \in K_1$ and every point $y \in K_2$ are separated by at least one of the intervals (c_i, d_i) . To be more precise one can integrate the inequality

$$\sum_{i=1}^{k} \left(\mathbf{1}_{[a,c_i]}(x) \mathbf{1}_{[d_i,b]}(y) + \mathbf{1}_{[a,c_i]}(y) \mathbf{1}_{[d_i,b]}(x) \right) \ge \mathbf{1}_{K_1}(x) \mathbf{1}_{K_2}(y) + \mathbf{1}_{K_1}(y) \mathbf{1}_{K_2}(x).$$

against $e^x e^y dx dy$ and use Fubini.

5. Log-BM inequality

In this chapter we shall need the following definition.

Definition 4.

- (1) We say that a function $f : \mathbb{R}^n \to \mathbb{R}$ is unconditional if for any choice of signs $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, 1\}$ and any $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ we have $f(\varepsilon_1 x_1, \ldots, \varepsilon_n x_n) = f(x)$.
- (2) We say that an unconditional function is decreasing if for any $1 \le i \le n$ and any real numbers $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ the function

$$t \mapsto f(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n)$$

is non-increasing on $[0, \infty)$.

- (3) A set $A \subseteq \mathbb{R}^n$ is called an *ideal* if $\mathbf{1}_A$ is unconditional and decreasing. In other words, a set $A \subset \mathbb{R}^n$ is an ideal if $(x_1, \ldots, x_n) \in A$ implies $(\delta_1 x_1, \ldots, \delta_n x_n) \in A$ for any choice of $\delta_1, \ldots, \delta_n \in [-1, 1]$. In other words, an ideal is a union of symmetric coordinate boxes. The class of all ideals (in \mathbb{R}^n) will be denoted by \mathcal{K}_I . Note that $A, B \in \mathcal{K}_I$ implies $\lambda A + (1 - \lambda)B \in \mathcal{K}_I$.
- (4) A set $A \subseteq \mathbb{R}^n$ is called symmetric if A = -A. The class of all symmetric convex sets in \mathbb{R}^n will be denoted by \mathcal{K}_S .

(5) A measure μ on \mathbb{R}^n is called unconditional if it has an unconditional density with respect to the Lebesgue measure.

Definition 5. We say that a Borel measure μ on \mathbb{R}^n satisfies the Brunn-Minkowski inequality in the class of sets \mathcal{K} if for any $A, B \in \mathcal{K}$ and for any $\lambda \in [0, 1]$ we have

(29)
$$\mu(\lambda A + (1-\lambda)B)^{1/n} \ge \lambda \mu(A)^{1/n} + (1-\lambda)\mu(B)^{1/n}$$

Definition 6. Let \mathcal{K} be a class of subsets closed under dilations. We say that a family $\odot = (\odot_{\lambda})_{\lambda \in [0,1]}$ of functions $\mathcal{K} \times \mathcal{K} \to \mathcal{K}$ is a geometric mean if for any $A, B \in \mathcal{K}$ the set $A \odot_{\lambda} B$ is measurable, satisfies an inclusion $A \odot_{\lambda} B \subseteq \lambda A + (1 - \lambda)B$, and $(sA) \odot_{\lambda} (tB) = s^{\lambda}t^{1-\lambda}(A \odot_{\lambda} B)$, for any s, t > 0.

Definition 7. We say that a Borel measure μ on \mathbb{R}^n satisfies the log-Brunn-Minkowski inequality in the class of sets \mathcal{K} with a geometric mean \odot , if for any sets $A, B \in \mathcal{K}$ and for any $\lambda \in [0, 1]$ we have

$$\mu(A \odot_{\lambda} B) \ge \mu(A)^{\lambda} \mu(B)^{1-\lambda}.$$

Remark 8. We shall use two different geometric means. The first one is the geometric mean $\bigcirc^S : \mathcal{K}_S \times \mathcal{K}_S \to \mathcal{K}_S$, defined by the formula

$$A \odot^S_{\lambda} B = \{ x \in \mathbb{R}^n : \langle x, u \rangle \le h^{\lambda}_A(u) h^{1-\lambda}_B(u), \ \forall u \in S^{n-1} \}.$$

Here h_A is the support function of A, i.e., $h_A(u) = \sup_{x \in A} \langle x, u \rangle$.

The second mean $\bigcirc^I : \mathcal{K}_I \times \mathcal{K}_I \to \mathcal{K}_I$ is defined by

$$A \odot^{I}_{\lambda} B = \bigcup_{x \in A, y \in B} [-|x_{1}|^{\lambda} |y_{1}|^{1-\lambda}, |x_{1}|^{\lambda} |y_{1}|^{1-\lambda}] \times \ldots \times [-|x_{n}|^{\lambda} |y_{n}|^{1-\lambda}, |x_{n}|^{\lambda} |y_{n}|^{1-\lambda}].$$

It is straightforward to check, with the help of the inequality $a^{\lambda}b^{1-\lambda} \leq \lambda a + (1-\lambda)b$, $a, b \geq 0$, that both means are indeed geometric.

We have the following theorem.

Theorem 51. The log-Brunn-Minkowski inequality holds true with the geometric mean \bigcirc^{I} for any measure with unconditional log-concave density in the class \mathcal{K}_{I} of all ideals in \mathbb{R}^{n} .

Proof. Let $A, B \in \mathcal{K}_I$ and let us take $f, g, m : [0, +\infty)^n \to [0, +\infty)$ given by $f = \mathbf{1}_{A \cap [0, +\infty)^n}$, $g = \mathbf{1}_{B \cap [0, +\infty)^n}$ and $m = \mathbf{1}_{(A \odot_{\lambda}^I B) \cap [0, +\infty)^n}$. Let φ be the unconditional log-concave density of μ . We define

$$F(x) = f(e^{x_1}, \dots, e^{x_n})\varphi(e^{x_1}, \dots, e^{x_n})e^{x_1 + \dots + x_n}, \quad G(x) = g(e^{x_1}, \dots, e^{x_n})\varphi(e^{x_1}, \dots, e^{x_n})e^{x_1 + \dots + x_n},$$
$$M(x) = m(e^{x_1}, \dots, e^{x_n})\varphi(e^{x_1}, \dots, e^{x_n})e^{x_1 + \dots + x_n}.$$

One can easily check, using the definition of \mathcal{K}_I and the definition of the geometric mean \odot^I_{λ} , as well as the inequalities

$$\varphi(e^{\lambda x_1 + (1-\lambda)y_1}, \dots, e^{\lambda x_n + (1-\lambda)y_n})$$

$$\geq \varphi(\lambda e^{x_1} + (1-\lambda)e^{y_1}, \dots, \lambda e^{x_n} + (1-\lambda)e^{y_n}) \geq \varphi(e^{x_1}, \dots, e^{x_n})^{\lambda} \varphi(e^{y_1}, \dots, e^{y_n})^{1-\lambda},$$

that the functions F, G, M satisfy the assumptions of the Prékopa-Leindler inequality. As a consequence, we get $\mu((A \odot^I_{\lambda} B) \cap [0, +\infty)^n) \ge \mu(A \cap [0, +\infty)^n)^{\lambda} \mu(B \cap [0, +\infty)^n)^{1-\lambda}$. The assertion follows from unconditionality of our measure μ and the fact that A, B and $A \odot^I_{\lambda} B$ are ideals.

We shall prove that the log-BM inequality implies the BM-inequality.

Proposition 52. Suppose that a Borel measure μ with a radially decreasing density f, i.e. density satisfying $f(tx) \geq f(x)$ for any $x \in \mathbb{R}^n$ and $t \in [0, 1]$, satisfies the log-Brunn-Minkowski inequality, with a geometric mean \odot , in a certain class of sets \mathcal{K} . Then μ satisfies the Brunn-Minkowski inequality in the class \mathcal{K} .

Proof. Let us first assume that $\mu(A)\mu(B) > 0$. From the definition of geometric mean we have $A \odot_p B \subseteq pA + (1-p)B$, for any $p \in (0, 1)$. Thus,

$$\mu(\lambda A + (1-\lambda)B) = \mu\left(p \cdot \frac{\lambda}{p}A + (1-p) \cdot \frac{1-\lambda}{1-p}B\right) \ge \mu\left(\left(\frac{\lambda}{p}A\right) \odot_p\left(\frac{1-\lambda}{1-p}B\right)\right)$$
$$= \mu\left(\left(\frac{\lambda}{p}\right)^p \left(\frac{1-\lambda}{1-p}\right)^{1-p}A \odot_p B\right).$$

Let $t = \left(\frac{\lambda}{p}\right)^p \left(\frac{1-\lambda}{1-p}\right)^{1-p}$ and $C = A \odot_p B$. From the concavity of the logarithm it follows that $0 \le t \le 1$. We have

(30)
$$\mu(tC) = \int_{tC} f(x) \, \mathrm{d}x = t^n \int_C f(tx) \, \mathrm{d}x \ge t^n \int_C f(x) \, \mathrm{d}x = t^n \mu(C).$$

Therefore, since μ satisfies the log-Brunn-Minkowski inequality,

$$\mu(\lambda A + (1-\lambda)B) \ge t^n \mu(A \odot_p B) \ge t^n \mu(A)^p \mu(B)^{1-p} = \left[\left(\frac{\lambda}{p}\right)^p \left(\frac{1-\lambda}{1-p}\right)^{1-p} \right]^n \mu(A)^p \mu(B)^{1-p}.$$

Taking

(31)
$$p = \frac{\lambda \mu(A)^{1/n}}{\lambda \mu(A)^{1/n} + (1 - \lambda)\mu(B)^{1/n}}$$

gives

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

If, say, $\mu(B) = 0$ then by (30), applied for C replaced with A, and the fact that $0 \in B$ we get

$$\mu(\lambda A + (1-\lambda)B)^{1/n} \ge \mu(\lambda A)^{1/n} \ge \lambda \mu(A)^{1/n} = \lambda \mu(A)^{1/n} + (1-\lambda)\mu(B)^{1/n}.$$

As a consequence, applying our Proposition 52 we deduce the following theorem.

Theorem 53. Let μ be an unconditional log-concave measure on \mathbb{R}^n . Then μ satisfies the Brunn-Minkowski inequality in the class \mathcal{K}_I of all ideals in \mathbb{R}^n .

References

- [1] L. P. Arguin, Extrema of log-correlated random variables: Principles and Examples, arXiv:1601.00582
- [2] S. Artstein-Avidan, A. Giannopoulos, and V. D. Milman. Asymptotic geometric analysis. Part I, volume 202 of Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2015.
- [3] K. Ball. Cube slicing in \mathbb{R}^n , Proc. Amer. Math. Soc., 97(3):465–473, 1986.
- [4] F. Barthe, On a reverse form of the Brascamp-Lieb inequality, arXiv:math/9705210
- [5] F. Barthe, Optimal Young's inequality and its converse: a simple proof, arXiv:math/9704210
- [6] F. Barthe, O. Guédon, S. Mendelson, and A. Naor, A probabilistic approach to the geometry of the l_n^p -ball, Ann. Probab., 33(2):480–513, 2005.
- [7] C. Borell, The Ehrhard inequality, C. R. Math. Acad. Sci. Paris 337 (2003), no. 10, 663–666.

- [8] Cordero-Erausquin D., Fradelizi M., Maurey B., The (B) conjecture for the Gaussian measure of dilates of symmetric convex sets and related problems, J. Funct. Anal. 214 (2004), no. 2, 410–427.
- [9] Dembo, A., Cover, T. M., Thomas, J. A., Information-theoretic inequalities, IEEE Trans. Inform. Theory 37 (1991), no. 6, 1501–1518.
- [10] A. Eskenazis, P. Nayar, T. Tkocz, Gaussian mixtures: entropy and geometric inequalities, arXiv:1611.04921
- [11] R. Gardner, The Brunn-Minkowski inequality, Bull. Amer. Math. Soc. 39 (2002), 355–405.
- [12] R. Latała and D. Matlak, Royen's proof of the gaussian correlation inequality, arXiv:1512.08776
- [13] D. A. Levin, Y. Peres, E. L. Wilmer, *Markov Chains and Mixing Times*, book available at http://yuvalperes.com/markovmixing.pdf
- [14] G. Livshyts, A. Marsiglietti, P. Nayar, A. Zvavitch, On the Brunn-Minkowski inequality for general measures with applications to new isoperimetric-type inequalities, arXiv:1504.04878
- [15] T. Royen, A simple proof of the Gaussian correlation conjecture extended to some multivariate gamma distributions, Far East J. Theor. Stat., 48(2):139–145, 2014.
- [16] O. Zeitouni, Branching random walks and Gaussian fields, lecture notes