AMSI SUMMER RESEARCH SCHOLARSHIPS 2025–26

Get a taste for Research this Summer

Inequalities in Geometry and Analysis

Vinnie Ng

Supervised by A/Prof. Jiakun Liu

The University of Sydney

Jane Street

AMSI AUSTRALIAN MATHEMATICAL SCIENCES INSTITUTE

Abstract

Inequalities are an essential tool in geometry and analysis. This paper surveys the role of the Brunn-Minkowski inequality in unifying geometric results, such as the isoperimetric inequality, and its functional generalisations, such as the inequalities of Prékopa-Leindler and Brascamp-Lieb. A particular emphasis is placed on the application of optimal transport to efficient proofs of these geometric and functional inequalities.

1 Introduction

Inequalities are an essential tool in analysis and geometry. For example, they are central to a priori estimates which can prove the existence, uniqueness, and regularity of solutions to partial differential equations. Inequalities also capture fundamental relationships between geometric quantities such as surface area and volume. Thus, understanding the established inequalities and their connections is crucial for studying analysis and geometry within a modern unified framework.

Gardner (2002) extensively surveyed the role of the Brunn-Minkowski inequality in unifying geometry and analysis. While the Brunn-Minkowski inequality is concerned with the volume of sums of sets, it acts as a bridge between geometric properties of convex bodies implied by it and the broad analytic results obtained by generalising it to functions. A key technique used in the proof of the Brunn-Minkowski inequalities and its relatives is mass transport, which involves finding a map which transforms one distribution into another with equal total mass. Mass transport maps which minimise some cost functional are known as optimal transport maps, and have emerged as an efficient tool for proving inequalities in geometry and analysis.

The aim of this paper is to survey the Brunn-Minkowski inequality and its relatives in geometry and analysis, with a particular emphasis on their proofs by optimal transport methods. Figure 1 illustrates the scope of this paper and the relationships between the inequalities that will be discussed.

2 Statement of Authorship

The inequalities and proofs discussed in this paper are established results in the fields of geometry and analysis and have been cited accordingly. However, the exposition of optimal transport as a connecting methodology and the reframing of these results within a consistent framework and notation is my own work.

3 Basic Definitions and Notation

The sum $X + Y$ of two sets X and Y refers to the *Minkowski sum*, defined by $X + Y := \{x + y : x \in X, y \in Y\}$. The scalar multiple cX of a scalar $c \in \mathbb{R}$ and set X is defined by $cX = \{cx : x \in X\}$.

A *convex body* is a compact, convex subset of $\mathbb{R}^n$ with non-empty interior. The closed unit ball centred at the origin is an example of a convex body and is denoted B. The volume of a convex body or a compact domain

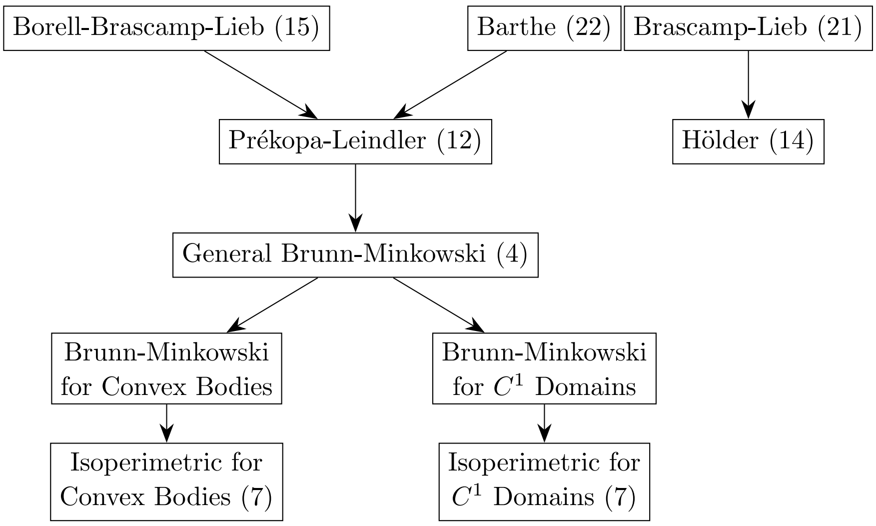

Figure 1: Relationships between the inequalities surveyed in this paper.

with C^1 boundary, denoted by V, is given by the n-dimensional Lebesgue measure and the surface area is given by the Minkowski definition.

$$S(K) = \lim_{\epsilon \to 0^+} \frac{V(K + \epsilon B) - V(K)}{\epsilon} \tag{1}$$

Two sets are *homothetic* if they differ only by translation and/or dilation.

A function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if, for all $x, y \in \mathbb{R}^n$ and $0 < \lambda < 1$, $f((1-\lambda)x + \lambda y) \le (1-\lambda)f(x) + \lambda f(y)$. The function f is *concave* if $-f$ is convex.

4 p-Means and p-Concave Functions

For finite sets, there are several types of means, including the arithmetic mean (AM) and the geometric mean (GM), which are related by the well known AM-GM inequality $(\prod_{i=1}^n a_i)^{\frac{1}{n}} \le \frac{1}{n}\sum_{i=1}^n a_i$, where $a_1, \ldots, a_n$ are positive real numbers. These means are special cases of p-means, which both generalise the notion of the mean and allow for weighted analogues. The following definition follows the convention used in Hardy et al. (1934).

Definition 4.1 (p-mean)**.** Let $\mathbf{a}, \boldsymbol{\lambda} \in \mathbb{R}^n$ such that $a_i \ge 0$ and $0 < \lambda_i < 1$ for all $1 \le i \le n$ and $\sum_{i=1}^n \lambda_i = 1$. If $p \ne 0$, the p-mean of the numbers $\mathbf{a}$ with weights $\boldsymbol{\lambda}$ is defined by

$$M_p(\mathbf{a}, \boldsymbol{\lambda}) = \left(\sum_{i=1}^n \lambda_i a_i^p \right)^{\frac{1}{p}} \tag{2}$$

except when both $p < 0$ and $\prod_{i=1}^n a_i = 0$, in which case $M_p(\mathbf{a}, \boldsymbol{\lambda}) = 0$. If $p = 0$, the p-mean is defined by $M_0(\mathbf{a}, \boldsymbol{\lambda}) = \prod_{i=1}^n a_i^{\lambda_i}$. If $p = \pm\infty$, the p-mean is defined by $M_\infty(\mathbf{a}, \boldsymbol{\lambda}) = \max_{1 \le i \le n} a_i$ and $M_{-\infty}(\mathbf{a}, \boldsymbol{\lambda}) = \min_{1 \le i \le n} a_i$ respectively.

The notation $M_p(a_1, a_2; \lambda) := M_p((a_1, a_2), (1-\lambda, \lambda))$ will be used in the $n = 2$ case.

Note that the AM and GM are the special cases of the p-means with equal weights and $p = 1$ and $p = 0$ respectively. The definition of the p-means when $p = 0$ and $p = \pm\infty$ is consistent with the limiting behaviour of the definition when $p \neq 0$, as detailed in the Appendix (Section 13.1). The p-means satisfy Jensen's inequality, a generalisation of the AM-GM inequality to the p-means (Hardy et al. 1934), which is proven in the Appendix (Section 13.1).

Theorem 4.1 (Jensen's inequality for means)**.** Let $\mathbf{a}, \boldsymbol{\lambda} \in \mathbb{R}^n$ satisfy the conditions in Definition 4.1. If $p < q$,

$$M_p(\mathbf{a}, \boldsymbol{\lambda}) < M_q(\mathbf{a}, \boldsymbol{\lambda}) \tag{3}$$

The p-means allow the definition of a concave function to be generalised by interpolating using weighted p-means rather than using linear interpolation. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *p-concave* if, for all $x, y \in \mathbb{R}^n$ and $0 < \lambda < 1$, $f((1-\lambda)x + \lambda y) \geq M_p(f(x), f(y); \lambda)$. If $p > 0$, f is p-concave if and only if f^p is concave, and if $p = 0$, f is p-concave if and only if $\log f$ is concave (also known as log-concave). Moreover, by Jensen's inequality (Theorem 4.1), if f is p-concave, then f is also q-concave for all $q < p$.

5 The Brunn-Minkowski Inequality

The Brunn-Minkowski inequality relates the volume of sum sets to the volume of the individual sets. Its measure theoretic formulation is known as the general Brunn-Minkowski inequality, and it acts as a critical link between the complementary fields of geometry and analysis. In the words of Gardner (2002), "the Brunn-Minkowski inequality appears like an octopus, tentacles reaching far and wide, its shape and colour changing as it roams from one area to the next".

Theorem 5.1 (General Brunn-Minkowski inequality)**.** Let $0 < \lambda < 1$ and let X and Y be non-empty, bounded, measurable subsets of $\mathbb{R}^n$ such that $(1-\lambda)X + \lambda Y$ is measurable. Then,

$$V((1-\lambda)X + \lambda Y)^{\frac{1}{n}} \geq (1-\lambda)V(X)^{\frac{1}{n}} + \lambda V(Y)^{\frac{1}{n}} \tag{4}$$

Equality holds if and only if there exist homothetic convex bodies X^* and Y^* which differ from X and Y respectively by sets of measure zero.

Gardner (2002) presents two equivalent forms of the Brunn-Minkowski inequality. One such form is

$$V(X+Y)^{\frac{1}{n}} \geq V(X)^{\frac{1}{n}} + V(Y)^{\frac{1}{n}} \tag{5}$$

which is the specific case of (4) when $\lambda = \frac{1}{2}$. (4) can be recovered by replacing X and Y with $(1-\lambda)X$ and λY respectively and applying the homogeneity of volume. Another equivalent inequality is

$$V((1-\lambda)X + \lambda Y) \geq \min\{V(X), V(Y)\} \tag{6}$$

which follows from (4) by Jensen's inequality (Theorem 4.1). By assuming without loss of generality that $V(X)V(Y) > 0$, (4) can be recovered by replacing X and Y with $V(X)^{-\frac{1}{n}}X$ and $V(Y)^{-\frac{1}{n}}Y$ respectively and

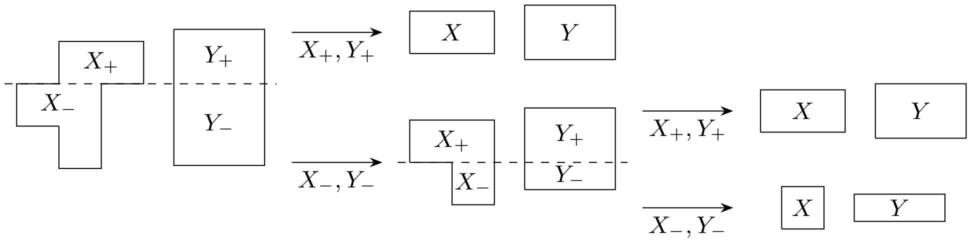

Figure 2: Hadwiger-Ohmann cuts for two sets X and Y composed of finite unions of boxes in $\mathbb{R}^2$

substituting

$$\lambda = \frac{V(Y)^{\frac{1}{n}}}{V(X)^{\frac{1}{n}} + V(Y)^{\frac{1}{n}}}$$

This yields (5), which is equivalent to (4).

Hadwiger and Ohmann (1956) proved (5) by considering the case where X and Y are finite unions of boxes and inductively cutting the domains into simpler subdomains on which (5) holds. The result was extended to the general case via an approximation argument.

Proof of Theorem 5.1 using Hadwiger-Ohmann cuts. Consider the case where $X, Y \subset \mathbb{R}^n$ are boxes with side lengths $(x_i)_{i=1}^n$ and $(y_i)_{i=1}^n$ respectively. Then, $X + Y$ is a box with side lengths $(x_i + y_i)_{i=1}^n$. By the AM-GM inequality,

$$\frac{V(X)^{\frac{1}{n}} + V(Y)^{\frac{1}{n}}}{V(X+Y)^{\frac{1}{n}}} = \left(\prod_{i=1}^n \frac{x_i}{x_i + y_i}\right)^{\frac{1}{n}} + \left(\prod_{i=1}^n \frac{y_i}{x_i + y_i}\right)^{\frac{1}{n}} \leq \frac{1}{n}\sum_{i=1}^n \frac{x_i}{x_i + y_i} + \frac{1}{n}\sum_{i=1}^n \frac{y_i}{x_i + y_i} = 1$$

Now, consider the case where X and Y are finite unions of boxes. Let $N(X)$ $(N(Y))$ denote the minimal number of boxes in a decomposition of X (Y). The proof proceeds by induction on $N(X)+N(Y)$. The base case where $N(X) + N(Y) = 2$ has already been proven, and for $N(X) + N(Y) = m > 2$, the induction hypothesis is that (5) holds when $2 \leq N(X) + N(Y) \leq m - 1$. The key step in the inductive case is the Hadwiger-Ohmann cut. By translating if necessary, X and Y can be partitioned by a hyperplane $\{x_k = 0\}$ for some $1 \leq k \leq n$ such that $\frac{V(X_\pm)}{V(X)} = \frac{V(Y_\pm)}{V(Y)}$, where $X_+ = X \cap \{x_k \geq 0\}$, $X_- = X \cap \{x_k \leq 0\}$, and Y_+ and Y_- are defined analogously. Moreover, it is always possible to choose a cut that ensures that either $N(X_+) < N(X)$ and $N(X_-) < N(X)$ or the equivalent condition for Y, so it follows that $N(X_\pm) + N(Y_\pm) < N(X) + N(Y)$. Figure 2 illustrates an example of this process in the $n = 2$ case.

Note that $X_+ + Y_+ \subset \{x_k \geq 0\}$ and $X_- + Y_- \subset \{x_k \leq 0\}$, so they are almost disjoint subsets of $X \cup Y$. This implies that $V(X + Y) \geq V(X_+ \cup Y_+) + V(X_- \cup Y_-)$. By the induction hypothesis and the assumption that $\frac{V(X_\pm)}{V(X)} = \frac{V(Y_\pm)}{V(Y)}$, it follows that

$$\begin{aligned} V(X+Y) &\geq \left(V(X_+)^{\frac{1}{n}} + V(Y_+)^{\frac{1}{n}}\right)^n + \left(V(X_-)^{\frac{1}{n}} + V(Y_-)^{\frac{1}{n}}\right)^n \\ &= V(X_+)\left(1 + \frac{V(Y)}{V(X)}^{\frac{1}{n}}\right)^n + V(X_-)\left(1 + \frac{V(Y)}{V(X)}^{\frac{1}{n}}\right)^n \\ &= V(X)\left(1 + \frac{V(Y)}{V(X)}^{\frac{1}{n}}\right)^n = V(X)^{\frac{1}{n}} + V(Y)^{\frac{1}{n}} \end{aligned}$$

To obtain the weaker conditions in Theorem 5.1, extend (5) to compact sets by approximating them from the outside by finite unions of boxes (Lemma 13.1) and then to measurable sets by approximating them from the inside by compact sets (Lemma 13.2). The details of the argument are provided in the Appendix (Section 13.2). □

The key idea behind the Hadwiger-Ohmann cut is that it partitions the sets such that, after volume normalisation, the volume on each side of the cut is the same for both sets. Hence, each cut can be thought of as a one-dimensional mass transport map between the two sets which enables the two sets to be more easily compared after reducing them to the base case.

The Brunn-Minkowski inequality also holds if X and Y are both convex bodies or compact domains with C^1 boundary (Gardner 2002). In these cases, $(1-\lambda)X + \lambda Y$ is automatically measurable because X and Y are both compact (Lemma 13.1). Moreover, the Brunn-Minkowski inequality for convex bodies has a geometric interpretation (Gardner 2002). Let $K \subset \mathbb{R}^{n+1}$ be a convex body and let $u \in S^n$. Let $u^\perp = \{x \in \mathbb{R}^{n+1} : \langle x, u \rangle = 0\}$ denote the n-dimensional hyperplane that passes through the origin and is orthogonal to the vector u. The geometric interpretation is that the function $f(t) = V_n(K \cap (u^\perp + tu))$, which gives the n-dimensional volume of the parallel hyperplane sections of K, is $\frac{1}{n}$-concave.

6 The Isoperimetric Inequality

Isoperimetric inequalities are geometric inequalities that relate the surface area and volume of geometrical figures. The simplest version is the classical isoperimetric inequality for closed curves on the plane, which bounds the length of the curve L from below by the enclosed area A according to $L^2 \geq 4\pi A$. Geometrically, this means that the circle encloses the maximum possible area for a given perimeter. There exist analogous isoperimetric inequalities for convex bodies and compact domains with C^1 boundary in $\mathbb{R}^n$ (Gardner 2002).

Theorem 6.1 (Isoperimetric inequality)**.** Let $K \subset \mathbb{R}^n$ be a convex body or a compact domain with C^1 boundary. Then,

$$\left(\frac{V(K)}{V(B)}\right)^{\frac{1}{n}} \leq \left(\frac{S(K)}{S(B)}\right)^{\frac{1}{n-1}} \tag{7}$$

Equality holds if and only if K and B are homothetic.

The isoperimetric inequality is a natural consequence of the Brunn-Minkowski inequality.

Proof of Theorem 6.1 using the Brunn-Minkowski inequality. Let K be a convex body or compact domain with C^1 boundary. Fix $\epsilon > 0$ and apply the Brunn-Minkowski inequality and the homogeneity of volume.

$$V(K+\epsilon B) \geq \left(V(K)^{\frac{1}{n}} + V(\epsilon B)^{\frac{1}{n}}\right)^n = \left(V(K)^{\frac{1}{n}} + \epsilon V(B)^{\frac{1}{n}}\right)^n \tag{8}$$

Note that the surface area and volume of a ball can be related by spherical shell integration.

$$V(B) = \int_0^1 S(rB)\,dr = S(B)\int_0^1 r^{n-1}\,dr = \frac{S(B)}{n} \tag{9}$$

Compute the surface area of K using (1) and apply (8), L'Hôpital's Rule, and (9).

$$S(K) := \lim_{\epsilon \to 0^+} \frac{V(K + \epsilon B) - V(K)}{\epsilon} \geq \lim_{\epsilon \to 0^+} n \left(V(K)^{\frac{1}{n}} + \epsilon V(B)^{\frac{1}{n}} \right)^{n-1} V(B)^{\frac{1}{n}}$$
$$= nV(K)^{\frac{n-1}{n}} V(B)^{\frac{1}{n}} = S(B)V(K)^{\frac{n-1}{n}} V(B)^{\frac{1-n}{n}}$$

The equality condition follows from the equality condition for the Brunn-Minkowski inequality. □

7 Optimal Transport and its Applications to Geometric Inequalities

The optimal transport problem is concerned with finding a cost-minimising plan to transport a source distribution to a target distribution. The source and target distributions are represented by Radon probability measures μ^+ and μ^- respectively. The transport costs are modelled by a cost density $c : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$, $(x, y) \mapsto c(x, y)$ representing the cost of moving a unit mass from x to y. The problem was first formulated by Monge.

Definition 7.1 (Monge Formulation)**.** Let the class of transport maps be defined as the set

$$\mathcal{A} = \left\{ T : \mathbb{R}^n \to \mathbb{R}^n \;\middle|\; \int_{\mathbb{R}^n} h(T(x))\, d\mu^+(x) = \int_{\mathbb{R}^n} h(y)\, d\mu^-(y) \;\forall\, h \in C^0(\mathbb{R}^n) \right\}$$

The optimal transport map $T^* : \mathbb{R}^n \to \mathbb{R}^n$ satisfies $I[T^*] = \inf_{T \in \mathcal{A}} I[T]$, where I is the cost functional

$$I[T] = \int_{\mathbb{R}^n} c(x, T(x))\, d\mu^+(x)$$

Monge's formulation was ill-posed because it has no solution if the source or target distributions are Dirac point masses. Moreover, the nonlinear nature of the optimisation problem made it difficult to solve. Later, Kantorovich addressed these limitations through his relaxed formulation of the problem.

Definition 7.2 (Kantorovich Formulation)**.** Let the class of transport plans be defined as the set of measures

$$\mathcal{M} = \left\{ \text{Radon probability measures } \mu \text{ on } \mathbb{R}^n \times \mathbb{R}^n \mid \mu(E \times \mathbb{R}^n) = \mu^+(E),\ \mu(\mathbb{R}^n \times E) = \mu^-(E)\ \forall E \in \mathcal{B}(\mathbb{R}^n) \right\}$$

The optimal transport plan μ^* satisfies $J[\mu^*] = \inf_{\mu \in \mathcal{M}} J[\mu]$, where J is the relaxed cost functional

$$J[\mu] = \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x, y)\, d\mu(x, y)$$

Kantorovich's formulation was well-posed due to its measure theoretic nature. More importantly, Kantorovich introduced a dual variational principle which could be applied to solve the problem. The details of these formulations and Kantorovich's solution to the problem are discussed in Evans (2001).

Suppose the cost density is a quadratic $c(x, y) = \frac{1}{2}\|x - y\|^2$ and the measures admit non-negative probability density functions (i.e. $d\mu^\pm = f^\pm\, dx$). In this setting, the optimal transport map under the Monge formulation (Definition 7.1) is known as the Brenier map, which has the following property (Evans 2001; Barthe 1998).

Theorem 7.1. Let $f^+, f^- : \mathbb{R}^n \to [0, \infty)$ be probability density functions. There exists a convex potential $\phi : \mathbb{R}^n \to \mathbb{R}$ such that $T = D\phi$ is the optimal transport map under the Monge formulation. That is, for each $h \in C^0(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} h(T(x)) f^+(x)\, dx = \int_{\mathbb{R}^n} h(x) f^-(x)\, dx$$

The convex potential ϕ can be considered a generalised solution (in the Alexandrov sense) to the Monge-Ampère equation.

$$f^-(D\phi)\det(D^2\phi) = f^+$$

Cafarelli's regularity theory for these solutions provides sufficient conditions for the convex potential to be a classical solution to the Monge-Ampère equation (Evans 2001).

Theorem 7.2. Let $\Omega^\pm \subset \mathbb{R}^n$ be bounded, connected, open sets and let $f^+, f^- : \Omega^\pm \to (0, \infty)$ be probability density functions such that $f^\pm$ and $\frac{1}{f^\pm}$ are bounded.

1. (Interior Regularity) If Ω^- is convex and there exists some $0 < \beta < 1$ such that $f^+ \in C^\beta(\Omega^+)$ and $f^- \in C^\beta(\Omega^-)$, then $\phi \in C^{2,\alpha}_{\text{loc}}(\Omega^+)$ for all $0 < \alpha < \beta$.
2. (Boundary Regularity) If both Ω^+ and Ω^- are uniformly convex and $\partial\Omega^+$ and $\partial\Omega^-$ are C^∞, then $\phi \in C^{2,\alpha}(\overline{\Omega^+})$ for some $0 < \alpha < 1$.

The Brenier map can be used to directly prove the Brunn-Minkowski inequality (Theorem 5.1) for convex bodies (Alesker et al. 1999).

Proof of Theorem 5.1 for convex bodies using the Brenier map. Let $K, L \subset \mathbb{R}^n$ be convex bodies. Assume that $V(K)V(L) > 0$, otherwise the inequality holds trivially. Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be the convex potential such that $T : K \to L$, $x \mapsto D\phi(x)$ is the Brenier map transporting mass between the densities $f^+ = \frac{1}{V(K)}$ and $f^- = \frac{1}{V(L)}$ defined on K and L respectively. By Theorem 7.2, $\phi \in C^2_{\text{loc}}(K)$, so it solves the Monge-Ampère equation $\det D^2\phi(x) = \frac{V(L)}{V(K)}$ for all $x \in K$.

Compute the volume of $K + L$, note that $(\text{id} + T)(K) \subseteq K + L$, and use the substitution $z = (\text{id} + T)(x)$

$$V(K+L) = \int_{K+L} dz \geq \int_{(\text{id}+T)(K)} dz = \int_K \det(I + DT(x))\, dx = \int_K \det(I + D^2\phi(x))\, dx \tag{10}$$

By the Monge-Ampère equation, the Hessian matrix $D^2\phi(x)$ is positive semidefinite for all $x \in K$. Then, by the $\frac{1}{n}$-concavity of the determinant operator (Lemma 13.3) and the Monge-Ampère equation,

$$\det(I + D^2\phi(x)) \geq \left(1 + (\det D^2\phi(x))^{\frac{1}{n}}\right)^n = \left(1 + \left(\frac{V(L)}{V(K)}\right)^{\frac{1}{n}}\right)^n \tag{11}$$

Therefore, it follows that

$$V(K+L) \geq \int_K \det(I + D^2\phi(x))\, dx \geq \left(1 + \left(\frac{V(L)}{V(K)}\right)^{\frac{1}{n}}\right)^n \int_K dx = \left(V(K)^{\frac{1}{n}} + V(L)^{\frac{1}{n}}\right)^n$$

The equality conditions can be recovered by a close examination of the proof. The inequality in (11) achieves equality when $D^2\phi$ is a positive multiple of the identity, which occurs if and only if $T = D\phi$ is a homothetic transformation. The inequality in (10) arises from the subset relationship $(\text{id} + T)(K) \subseteq K + L$, which achieves equality when T is a homothetic transformation. To see this, let $T(x) = cx + b$, where $c > 0$ and $b \in \mathbb{R}^n$, and $z \in K + L$. Then, there exists $x, y \in K$ such that

$$z = x + T(y) = x + cy + b = (1 + c)((1 - \lambda)x + \lambda y) + b = (\text{id} + T)((1 - \lambda)x + \lambda y)$$

when $\lambda = \frac{c}{1+c}$. Since K is convex, $(1-\lambda)x + \lambda y \in K$, so it follows that $K + L \subseteq (\mathrm{id} + T)(K)$. □

In this proof, the Brenier map plays a similar role to Hadwiger-Ohmann cuts in relating the volumes of the two convex bodies by transporting mass from one to the other. Additionally, the minimality of the Brenier map enables the Monge-Ampère equation to be applied, which gives rise to the inequality through the $\frac{1}{n}$-concavity of the determinant operator, a property which holds locally.

The Brenier map also offers a direct proof of the isoperimetric inequality (Theorem 6.1) for convex bodies with C^1 boundary (Guillen and McCann 2010).

Proof of Theorem 6.1 for convex bodies with C^1 boundary using the Brenier map. Let K be a convex body and let $r := \left(\frac{V(K)}{V(B)}\right)^{\frac{1}{n}}$, so that $V(K) = V(rB)$. Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be the convex potential such that $T : X \to Y$, $x \mapsto D\phi(x)$ is the Brenier map transporting mass between the densities $f^+ = 1$ and $f^- = 1$ defined on K and rB respectively. By Theorem 7.2, $\phi \in C^2(\overline{K})$, so it solves the Monge-Ampère equation $\det D^2\phi(x) = 1$ for all $x \in \overline{K} = K$.

Compute the volume of K and apply the Monge-Ampère equation and Lemma 13.4.

$$V(K) = \int_K dx = \int_K \left(\det D^2\phi(x)\right)^{\frac{1}{n}} dx \le \frac{1}{n}\int_K \operatorname{tr} D^2\phi(x)\, dx = \frac{1}{n}\int_K \Delta\phi(x)\, dx$$

Apply the divergence theorem, with ν denoting the outer unit normal vector, and the Cauchy-Schwarz inequality.

$$\int_K \Delta\phi(x)\, dx = \int_{\partial K} \langle D\phi(x), \nu(x)\rangle\, d\sigma(x) \le \int_{\partial K} \|D\phi(x)\|\, d\sigma(x)$$

Since $T(\partial K) \subset T(K) = rB$, for all $x \in \partial K$, $\|D\phi(x)\| = \|T(x)\| \le r$. Thus, it follows that

$$V(K) \le \frac{r}{n}\int_{\partial K} d\sigma(x) = \frac{1}{n}\left(\frac{V(K)}{V(B)}\right)^{\frac{1}{n}} S(K)$$

which is equivalent to the isoperimetric inequality, as shown in the proof of Theorem 6.1 using the Brunn-Minkowski inequality.

The equality conditions can be recovered by a similar method to in the proof of Theorem 5.1 using the Brenier map. □

In this proof, the geometries of the convex body and ball are able to be compared due to the application of the Brenier map to transport mass between them. Additionally, the surface area is recovered by the application of the divergence theorem to yield a surface integral.

8 The Prékopa-Leindler Inequality and the Borell-Brascamp-Lieb Inequality

Suppose f, g, and h are three non-negative, integrable functions and suppose that h dominates the functions f and g in the sense that $h((1-\lambda)x + \lambda y) \ge (1-\lambda)f(x) + \lambda g(y)$. By the linearity of the integral, the L^1-norm of h is at least the arithmetic mean of the L^1-norms of f and g. The Prékopa-Leindler inequality generalises this idea to the case where $h((1-\lambda)x+\lambda y) \geq f(x)^{1-\lambda}g(y)^{\lambda}$, which, by the AM-GM inequality, is a weaker assumption on the function h. In essence, it allows a pointwise estimate of h to be elevated to a global estimate of $\|h\|_1$.

Theorem 8.1 (Prékopa-Leindler inequality)**.** Let $0 < \lambda < 1$ and let $f, g, h : \mathbb{R}^n \to [0, \infty)$ be integrable functions. Suppose $h((1-\lambda)x+\lambda y) \geq f(x)^{1-\lambda}g(y)^{\lambda}$ for all $x, y \in \mathbb{R}^n$. Then,

$$\int_{\mathbb{R}^n} h(x)\,dx \geq \left(\int_{\mathbb{R}^n} f(x)\,dx\right)^{1-\lambda}\left(\int_{\mathbb{R}^n} g(x)\,dx\right)^{\lambda} \tag{12}$$

The Prékopa-Leindler inequality implies the general Brunn-Minkowski inequality when considering the volumes as integrals of indicator functions. Thus, the Prékopa-Leindler inequality is considered a functional generalisation of the general Brunn-Minkowski inequality.

Proof of Theorem 5.1 using the Prékopa-Leindler inequality. Let X and Y satisfy the hypotheses of Theorem 5.1 and let $f = \mathbb{1}_X$, $g = \mathbb{1}_Y$ and $h = \mathbb{1}_{(1-\lambda)X+\lambda Y}$. If $f(x)^{1-\lambda}g(y)^{\lambda} = 1$, then $x \in X$ and $y \in Y$, so it follows that $h((1-\lambda)x+\lambda y) = 1$. This means that the hypotheses of Theorem 8.1 hold, so, by the Prékopa-Leindler inequality and Jensen's inequality (Theorem 4.1),

$$V((1-\lambda)X+\lambda Y) = \int_{\mathbb{R}^n} h(x)\,dx \geq \left(\int_{\mathbb{R}^n} f(x)\,dx\right)^{1-\lambda}\left(\int_{\mathbb{R}^n} g(x)\,dx\right)^{\lambda} = V(X)^{1-\lambda}V(Y)^{\lambda} \geq \min\{V(X), V(Y)\}$$

which is (6), an equivalent form of the general Brunn-Minkowski inequality. □

Another form of the Prékopa-Leindler inequality can be obtained by choosing h as the smallest function that satisfies the inequality $h((1-\lambda)x+\lambda y) \geq f(x)^{1-\lambda}g(y)^{\lambda}$. This leads to the inequality

$$\overline{\int_{\mathbb{R}^n}} \sup\{f(x)^{1-\lambda}g(y)^{\lambda} : (1-\lambda)x+\lambda y = z\}\,dz \geq \left(\int_{\mathbb{R}^n} f(x)\,dx\right)^{1-\lambda}\left(\int_{\mathbb{R}^n} g(x)\,dx\right)^{\lambda} \tag{13}$$

where the upper Lebesgue integral is required since the supremum over the $(n-1)$-dimensional subspace need not be measurable. In this form, the Prékopa-Leindler inequality can be interpreted as a reverse form of Hölder's inequality, which is equivalent to

$$\int_{\mathbb{R}^n} f(x)^{1-\lambda}g(x)^{\lambda}\,dx \leq \left(\int_{\mathbb{R}^n} f(x)\,dx\right)^{1-\lambda}\left(\int_{\mathbb{R}^n} g(x)\,dx\right)^{\lambda} \tag{14}$$

The Prékopa-Leindler inequality is a special case of the Borell-Brascamp-Lieb inequality, which generalises the hypothesis by allowing f and g to be interpolated by a p-mean for $-\frac{1}{n} \leq p \leq \infty$.

Theorem 8.2 (Borell-Brascamp-Lieb inequality)**.** Let $0 < \lambda < 1$, $-\frac{1}{n} \leq p \leq \infty$, and $f, g, h : \mathbb{R}^n \to [0, \infty)$ be integrable functions. Suppose $h((1-\lambda)x+\lambda y) \geq M_p(f(x), g(y); \lambda)$ for all $x, y \in \mathbb{R}^n$. Then, with $\gamma = \frac{p}{np+1}$,

$$\int_{\mathbb{R}^n} h(x)\,dx \geq M_{\gamma}\left(\int_{\mathbb{R}^n} f(x)\,dx, \int_{\mathbb{R}^n} g(x)\,dx; \lambda\right) \tag{15}$$

Like the Prékopa-Leindler inequality, the Borell-Brascamp-Lieb inequality has an equivalent form.

$$\overline{\int_{\mathbb{R}^n}} \sup\{M_p(f(z_1), g(z_2); \lambda) : (1-\lambda)z_1+\lambda z_2 = x\}\,dx \geq M_{\gamma}\left(\int_{\mathbb{R}^n} f(x)\,dx, \int_{\mathbb{R}^n} g(x)\,dx; \lambda\right) \tag{16}$$

The Borell-Brascamp-Lieb inequality can be proved in the $n = 1$ case by applying a one-dimensional mass transport map and integrating a p-mean inequality. The $n \geq 2$ case can be proven by induction on the dimension. The Prékopa-Leindler inequality can be proved similarly, using the AM-GM inequality. These methods of proof are detailed in Gardner (2002).

McCann (1994) developed an alternative proof of the Borell-Brascamp-Lieb inequality in the case where $-\frac{1}{n} < p \leq \infty$ using optimal transport. A key tool is the displacement interpolant, which provides an interpolation between two probability measures which can be linked by a Brenier map.

Definition 8.1 (Displacement interpolant)**.** Let μ^+ and μ^- be probability measures on $\mathbb{R}^n$. Assume that there exists a convex potential $\phi : \mathbb{R}^n \to \mathbb{R}$ such that $T : \mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto D\phi(x)$ is the Brenier map transporting mass between μ^+ and μ^-. For $0 \leq t \leq 1$, the displacement interpolant ρ_t between μ^+ and μ^- is defined as the probability measure on $\mathbb{R}^n$ that satisfies

$$\int_{\mathbb{R}^n} \xi(x)\, d\rho_t(x) = \int_{\mathbb{R}^n} \xi((1-t)y + tD\phi(y))\, d\mu^+(y) \tag{17}$$

for every $\xi \in C^0(\mathbb{R}^n)$. In other words, ρ_t is the pushforward of μ^+ by the function $x \mapsto (1-t)x + tD\phi(x)$.

To aid the proof of Theorem 8.2, define the restricted function space $\mathcal{C}^L(\mathbb{R}^n)$ of functions by

$$\mathcal{C}^L(\mathbb{R}^n) = \{f \in L^1(\mathbb{R}^n) : \exists g \in C^{0,1}(\mathbb{R}^n),\ g > 0,\ r > 0,\ x \in \mathbb{R}^n \text{ such that } f = \mathbb{1}_{B_r(x)} g\}$$

These functions are the restriction of a positive Lipschitz function to an open Euclidean ball. Additionally, for any non-negative function f, define the notation $\Omega_f = f^{-1}((0,\infty))$.

Remark 8.3. Barthe (1998) introduced the subspace $\mathcal{C}^L(\mathbb{R}^n)$ and remarked that the Brenier map between any two functions in $\mathcal{C}^L(\mathbb{R}^n)$ will have sufficient regularity to satisfy the Monge-Ampère equation. To see this, consider $f^+, f^- \in \mathcal{C}^L(\mathbb{R}^n)$. By the definition of $\mathcal{C}^L(\mathbb{R}^n)$, Ω_{f^+} and Ω_{f^-} are open Euclidean balls and therefore convex. Moreover, since there exists some positive Lipschitz function g^+ such that $f^+ = \mathbb{1}_{\Omega_{f^+}} g^+$ and g^+ is bounded on the compact set $\overline{\Omega_{f^+}}$ because it is continuous, f^+ is bounded on Ω_{f^+}. Similarly, $\frac{1}{f^+}$ is bounded on Ω_{f^+} and f^- and $\frac{1}{f^-}$ are bounded on Ω_{f^-}. Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be the convex potential such that $T : \Omega_{f^+} \to \Omega_{f^-}$, $x \mapsto D\phi(x)$ is the Brenier map transporting mass between the densities f^+ and f^- on Ω_{f^+} and Ω_{f^-} respectively. By Theorem 7.2, $\phi \in C^2_{\text{loc}}(\Omega_{f^+})$, so it solves the Monge-Ampère equation for all $x \in \Omega_{f^+}$.

The displacement interpolant ρ_t admits a probability density function in the case where μ^+ and μ^- admit probability density functions; the details of the proof are provided in the Appendix (Section 13.4).

Lemma 8.4. Let $f^+, f^- \in \mathcal{C}^L(\mathbb{R}^n)$ be probability density functions such that $d\mu^\pm = f^\pm\, dx$. Define the function $s_t : \mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto (1-t)x + tD\phi(x)$. Then, the displacement interpolant ρ_t between μ^+ and μ^- admits a probability density function $\rho_t(x; f^+, f^-)$ which satisfies

$$\rho_t(x; f^+, f^-) = f^+(s_t^{-1}(x)) \det(Ds_t(s_t^{-1}(x)))^{-1} \tag{18}$$

for almost every x, with respect to the Lebesgue measure.

We first prove Theorem 8.2 for the restricted space of functions $\mathcal{C}^L(\mathbb{R}^n)$. The key idea in the proof is that the displacement interpolant ρ_t is dominated by the integrand of the left hand side of (16), denoted by

$$h_p(x; f, g, \lambda) := \sup\{M_p(f(z_1), g(z_2); \lambda) : (1-\lambda)z_1 + \lambda z_2 = x\} \tag{19}$$

Lemma 8.5. Let $f, g \in \mathcal{C}^L(\mathbb{R}^n)$. Then, for $p > -\frac{1}{n}$, (16) holds.

Proof. Assume without loss of generality that $\|f\|_1 \neq 0$ and $\|g\|_1 \neq 0$. Define the normalised functions $\tilde{f} = \frac{f}{\|f\|_1}$ and $\tilde{g} = \frac{g}{\|g\|_1}$. Let $S_k : L^1(\mathbb{R}^n) \to L^1(\mathbb{R}^n)$ be the volume-preserving scaling operator defined by $(S_k f)(x) := k^{-n} f\left(\frac{x}{k}\right)$. Define the constants $C := ((1-\lambda)\|f\|_1^\gamma + \lambda \|g\|_1^\gamma)^{\frac{1}{\gamma}}$ (the right hand side of (16)), $C_1 := \frac{C}{\|f\|_1}$ and $C_2 := \frac{C}{\|g\|_1}$.

Define $f^+ := S_{C_1^\gamma} \tilde{f}$ and $f^- := S_{C_2^\gamma} \tilde{g}$. Note that f^+ and f^- are probability density functions because they are the volume-preserving scaling of the functions $\tilde{f}$ and $\tilde{g}$ respectively which are both integrable and normalised. Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be the convex potential such that $T : \Omega_{f^+} \to \Omega_{f^-}$, $x \mapsto D\phi(x)$ is the Brenier map transporting mass between the densities f^+ and f^- defined on Ω_{f^+} and Ω_{f^-} respectively. By Remark 8.3, ϕ solves the Monge-Ampère equation.

Let $z \in \Omega_{f^+}$ and let $x := s_\lambda(z)$, where s_λ is as defined in the statement of Lemma 8.4. By (18), which holds for almost every $z \in \Omega_{f^+}$ with respect to the Lebesgue measure, and the $\frac{1}{n}$-concavity of the determinant operator (Lemma 11),

$$\rho_\lambda(x; f^+, f^-) = f^+(z) \left(\det((1-\lambda)I + \lambda D^2\phi(z))\right)^{-1} \leq f^+(z) \left((1-\lambda) + \lambda(\det D^2\phi(z))^{\frac{1}{n}}\right)^{-n}$$

By the Monge-Ampère equation,

$$\rho_\lambda(x; f^+, f^-) \leq f^+(z) \left((1-\lambda) + \lambda \left(\frac{f^+(z)}{f^-(D\phi(z))}\right)^{\frac{1}{n}}\right)^{-n} = \left((1-\lambda) f^+(z)^{-\frac{1}{n}} + \lambda f^-(D\phi(z))^{-\frac{1}{n}}\right)^{-n}$$

By Jensen's inequality (Theorem 4.1) and (19),

$$\rho_\lambda(x; f^+, f^-) \leq M_{-\frac{1}{n}}(f^+(z), f^-(D\phi(z)); \lambda) \leq M_p(f^+(z), f^-(D\phi(z)); \lambda) \leq h_p(x; f^+, f^-, \lambda) \tag{20}$$

for almost every $x \in s_\lambda(\Omega_{f^+})$.

By computing the integral of h_p and applying Lemma 13.5,

$$\int_{\mathbb{R}^n} h_p(x; f, g, \lambda)\, dx \geq \int_{s_\lambda(\Omega_{f^+})} h_p(x; f, g, \lambda)\, dx = C \int_{s_\lambda(\Omega_{f^+})} h_p\left(x; S_{C_1^\gamma}\tilde{f}, S_{C_2^\gamma}\tilde{g}; C_2^{-\gamma}\lambda\right) dx$$

By applying (20) and by noting that ρ_λ is the pushforward of the probability measure with density f^+ supported on Ω_{f^+},

$$\int_{\mathbb{R}^n} h_p(x; f, g, \lambda)\, dx \geq C \int_{s_\lambda(\Omega_{f^+})} \rho_\lambda(x; S_{C_1^\gamma} f, S_{C_2^\gamma} g)\, dx = C \int_{\Omega_{f^+}} dx = C := M_\gamma\left(\int_{\mathbb{R}^n} f(x)\, dx, \int_{\mathbb{R}^n} g(x)\, dx; \lambda\right)$$

□

Lemma 8.5 can be extended to the more general hypotheses of Theorem 8.2 when $-\frac{1}{n} < p \leq \infty$ by a series of approximation arguments. Note that the p-mean is monotonic in each input. By the monotone convergence theorem, (16) generalises to the case where f and g are positive Lipschitz functions on $\mathbb{R}^n$, since such functions can be written as a pointwise limit of an increasing sequence of functions in $\mathcal{C}^L(\mathbb{R}^n)$. Moreover, the monotone convergence theorem implies that (16) holds when f and g are simple functions of the form $\sum_{j=1}^k a_j \mathbb{1}_{K_j}$, where $a_j \in \mathbb{R}^n$ and $K_j \subset \mathbb{R}^n$ is compact, because such simple functions can be written as a pointwise limit of a decreasing sequence of positive Lipschitz functions. Theorem 8.2 follows by approximating the non-negative functions $f, g \in L^1(\mathbb{R}^n)$ as the pointwise limit of an increasing sequence of simple functions and applying the monotone convergence theorem. Note that McCann's proof does not generalise to the case where $p = -\frac{1}{n}$ because, in this limit, $\gamma \to -\infty$, so the scaling operation would no longer be well defined.

McCann's proof of the Borell-Brascamp-Lieb inequality shares many similarities with the proof of the Brunn-Minkowski inequality using the Brenier map. Both proofs apply the displacement interpolant to transport mass, but the proof of the Brunn-Minkowski inequality considers the equal weight interpolant $\rho_{\frac{1}{2}}$. Moreover, the inequality in both proofs arises from the $\frac{1}{n}$-concavity of the determinant operator.

9 The Brascamp-Lieb Inequality and Barthe's Inequality

The Brascamp-Lieb inequality and Barthe's inequality are far reaching analytic generalisations of the functional and geometric inequalities discussed so far (Gardner 2002). Their statement is as follows.

Theorem 9.1 (Brascamp-Lieb inequality and Barthe's inequality)**.** Let m, n be integers. For $i = 1, \dots, m$, let $c_i > 0$ and $n_i \in \mathbb{N}$ such that $\sum_{i=1}^m c_i n_i = n$, let $f_i \in L^1(\mathbb{R}^{n_i})$ be non-negative, and let $B_i : \mathbb{R}^n \to \mathbb{R}^{n_i}$ be a linear surjection. Assume that $\bigcap_{1\leq i\leq m} \ker B_i = \{0\}$. Then, the Brascamp-Lieb inequality states that

$$\int_{\mathbb{R}^n} \prod_{i=1}^m f_i(B_i x)^{c_i} \, dx \leq D^{-\frac{1}{2}} \prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} f_i(x) \, dx \right)^{c_i} \tag{21}$$

and Barthe's (reverse Brascamp-Lieb) inequality states that

$$\overline{\int_{\mathbb{R}^n}} \sup \left\{ \prod_{i=1}^m f_i(x_i)^{c_i} : x = \sum_{i=1}^m c_i B_i^* x_i,\ x_i \in \mathbb{R}^{n_i} \right\} dx \geq D^{\frac{1}{2}} \prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} f_i(x) \, dx \right)^{c_i} \tag{22}$$

where

$$D = \inf \left\{ \frac{\det\left(\sum_{i=1}^m c_i B_i^* A_i B_i\right)}{\prod_{i=1}^m (\det A_i)^{c_i}} : A_i \in \mathbb{R}^{n\times n} \text{ is symmetric and positive definite} \right\} \tag{23}$$

Note that the condition on $\bigcap_{1\leq i\leq m} \ker B_i$ simply ensures that $\sum_{i=1}^m c_i B_i^* A_i B_i$ is an isomorphism. If it is not satisfied, $D = 0$, so both inequalities hold trivially (Barthe 1998).

The Prékopa-Leindler inequality, in the form given in (13), is a corollary of Barthe's inequality.

Proof of Theorem 8.1 using Barthe's inequality. Let $m = 2$, $n_1 = n_2 = n$, and $B_1 = B_2 = \mathrm{id}$. Given $0 < \lambda < 1$, let $c_1 = 1 - \lambda$ and $c_2 = \lambda$. Let $f_1, f_2 \in L^1(\mathbb{R}^n)$ be non-negative. By Barthe's inequality,

$$\overline{\int_{\mathbb{R}^n}} \sup \left\{ f_1(z_1)^{1-\lambda} f_2(z_2)^{\lambda} : (1-\lambda) z_1 + \lambda z_2 = x \right\} dx \geq D^{\frac{1}{2}} \left(\int_{\mathbb{R}^n} f_1(x) \, dx \right)^{1-\lambda} \left(\int_{\mathbb{R}^n} f_2(x) \, dx \right)^{\lambda}$$

It remains to show that $D = 1$. Since $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ is a $\frac{1}{n}$-concave operator (Lemma 13.3), by Jensen's inequality (Theorem 4.1), it is also log-concave. Thus, for all $A_1, A_2 \in \mathbb{R}^{n\times n}$ and $0 < \lambda < 1$,

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

This inequality is sharp since it achieves equality when $A_1 = A_2$. Thus, it follows that $D = 1$. □

By an analogous proof, Hölder's inequality, in the form given in (14), is a corollary of the Brascamp-Lieb inequality, consistent with its interpretation as a reverse form of the Prékopa-Leindler inequality.

Theorem 9.1 will be proven using the Brenier map, a method due to Barthe (1998). Using the notation in Theorem 9.1, define the following:

$$I((f_i)_{i=1}^m) := \overline{\int_{\mathbb{R}^n}} \sup\left\{\prod_{i=1}^m f_i(x_i)^{c_i} : x = \sum_{i=1}^m c_i B_i^* x_i,\ x_i \in \mathbb{R}^{n_i}\right\} dx \qquad J((f_i)_{i=1}^m) := \int_{\mathbb{R}^n} \prod_{i=1}^m f_i(B_i x)^{c_i}\, dx$$

$$E := \inf\left\{\frac{I((f_i)_{i=1}^m)}{\prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} f_i(x)\, dx\right)^{c_i}} : f_i \in L^1(\mathbb{R}^{n_i}),\ f_i \geq 0\right\} = \inf\left\{I((f_i)_{i=1}^m) : f_i \in L^1(\mathbb{R}^{n_i}),\ \|f_i\|_1 = 1,\ f_i \geq 0\right\}$$

$$F := \sup\left\{\frac{J((f_i)_{i=1}^m)}{\prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} f_i(x)\, dx\right)^{c_i}} : f_i \in L^1(\mathbb{R}^{n_i}),\ f_i \geq 0\right\} = \sup\left\{J((f_i)_{i=1}^m) : f_i \in L^1(\mathbb{R}^{n_i}),\ \|f_i\|_1 = 1,\ f_i \geq 0\right\}$$

The proof is divided into several lemmas. First, consider the behaviour of I and J for *centred Gaussians*, functions of the form $G_A(x) = \exp(-\langle Ax, x\rangle)$ where A is a symmetric and positive definite matrix. Define E_g (F_g) analogously to E (F) but with the infimum (supremum) restricted to the space where each f_i is a centred Gaussian. These quantities can be directly computed; the details of the computations are provided in the Appendix (Section 13.5).

Lemma 9.2.

$$F_g := \sup\left\{\frac{J((G_{A_i})_{i=1}^m)}{\prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} G_{A_i}(x)\, dx\right)^{c_i}} : A_i \in \mathbb{R}^{n\times n} \text{ is symmetric and positive definite}\right\} = D^{-\frac{1}{2}} \tag{24}$$

Lemma 9.3.

$$E_g := \inf\left\{\frac{I((G_{A_i})_{i=1}^m)}{\prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} G_{A_i}(x)\, dx\right)^{c_i}} : A_i \in \mathbb{R}^{n\times n} \text{ is symmetric and positive definite}\right\} = D^{\frac{1}{2}} \tag{25}$$

The following inequality for functions in $\mathcal{C}^L(\mathbb{R}^n)$ is proven using Brenier maps. This differs from the previous proofs where only one Brenier map was required.

Lemma 9.4. For $1 \leq i \leq m$, let $f_i, h_i \in \mathcal{C}^L(\mathbb{R}^{n_i})$ be functions that satisfy $\int_{\mathbb{R}^{n_i}} f_i(x)\, dx = \int_{\mathbb{R}^{n_i}} h_i(x)\, dx$. Then,

$$I((f_i)_{i=1}^m) \geq DJ((h_i)_{i=1}^m) \tag{26}$$

Proof. Assume that $D > 0$, otherwise, the claim is trivial. By the homogeneity of I and J, it can be assumed that $\int_{\mathbb{R}^{n_i}} f_i(x)\, dx = \int_{\mathbb{R}^{n_i}} h_i(x)\, dx = 1$. For $1 \leq i \leq m$, let $\phi_i : \mathbb{R}^{n_i} \to \mathbb{R}$ be the convex potential such that $T_i : \Omega_{h_i} \to \Omega_{f_i}$, $x \mapsto D\phi_i(x)$ is the Brenier map transporting mass between the densities h_i and f_i on Ω_{h_i} and Ω_{f_i} respectively. By Remark 8.3, for all $x \in \Omega_{h_i}$, ϕ_i solves the Monge-Ampère equation

$$f_i(D\phi_i(x)) \det D^2\phi_i(x) = h_i(x) \tag{27}$$

Note that, for all $x \in \Omega_{h_i}$, $D^2\phi_i(x)$ is symmetric and positive semidefinite since ϕ_i is convex. Moreover, since f_i and h_i do not vanish on their respective domains, $\det D^2\phi_i(x) = \frac{h_i(x)}{f_i(D\phi_i(x))} \neq 0$ by (27). Thus, $D^2\phi_i(x)$ is symmetric and positive definite.

Define $S := \bigcap_{i=1}^m B_i^{-1}(\Omega_{h_i})$. Note that S is convex. To see this, let $x, y \in B_i^{-1}(\Omega_{h_i})$. Since B_i is linear and Ω_{h_i} is convex, $B_i((1-\lambda)x + \lambda y) = (1-\lambda)B_i x + \lambda B_i y \in \Omega_{h_i}$. By taking the preimage of B_i, it follows that $B_i^{-1}(\Omega_{h_i})$ is convex. Thus, S is convex since it is the intersection of convex sets.

Define $\Theta : S \to \mathbb{R}^n$ as $\Theta(y) := \sum_{i=1}^m c_i B_i^*(T_i(B_i y))$. For all $y \in S$, the Jacobian $D\Theta(y) = \sum_{i=1}^m c_i B_i^* DT_i(B_i y) B_i$ is symmetric since $DT_i(x) = D^2\phi_i(x)$ is symmetric for all $x \in \Omega_{h_i}$. Moreover, since $DT_i(x) = D^2\phi_i(x)$ is positive definite,

$$\det D\Theta(y) = \det\left(\sum_{i=1}^m c_i B_i^* DT_i(B_i y) B_i\right) \geq D \prod_{i=1}^m (\det DT_i(B_i y))^{c_i} > 0 \tag{28}$$

so $D\Theta(y)$ is positive definite.

We claim that Θ is injective. To see this, let $x, y \in S$, $x \neq y$ and suppose that $\Theta(x) = \Theta(y)$. Define $z : [0,1] \to \mathbb{R}^n$ by $z(t) = (1-t)x + ty$ and note that $z([0,1]) \in S$ since S is convex. Define $\psi : [0,1] \to \mathbb{R}$ by $\psi(t) := \langle \Theta(z(t)) - \Theta(x), y - x\rangle$. By the product rule,

$$\psi'(t) = \langle D\Theta(z(t))z'(t), y - x\rangle = \langle D\Theta(z(t))(y-x), y-x\rangle > 0 \tag{29}$$

since $D\Theta$ is positive definite on S. Since $\Theta(x) = \Theta(y)$, $\psi(0) = \psi(1) = 0$, so by Rolle's theorem, there exists $\tau \in (0,1)$ for which $\psi'(\tau) = 0$. This contradicts (29), so $x = y$.

By computing $J((h_i)_{i=1}^m)$, noting that the support of the integrand is S, and applying (27) and (28),

$$\begin{aligned} J((h_i)_{i=1}^m) = \int_S \prod_{i=1}^m h_i^{c_i}(B_i y)\,dy &= \int_S \prod_{i=1}^m (f_i(T_i(B_i y)) \det DT_i(B_i y))^{c_i}\,dy \\ &\leq \frac{1}{D}\int_S \left(\prod_{i=1}^m f_i(T_i(B_i y))^{c_i}\right) \det D\Theta(y)\,dy \\ &\leq \frac{1}{D}\overline{\int_S} \sup\left\{\prod_{i=1}^m f_i(x_i)^{c_i} : \Theta(y) = \sum_{i=1}^m c_i B_i^* x_i\right\} \det D\Theta(y)\,dy \end{aligned}$$

By substituting $x = \Theta(y)$, which is justified since Θ is a bijection onto its image,

$$\begin{aligned} J((h_i)_{i=1}^m) &\leq \frac{1}{D}\overline{\int_{\Theta(S)}} \sup\left\{\prod_{i=1}^m f_i(x_i)^{c_i} : x = \sum_{i=1}^m c_i B_i^* x_i\right\} dx \\ &\leq \frac{1}{D}\overline{\int_{\mathbb{R}^n}} \sup\left\{\prod_{i=1}^m f_i(x_i)^{c_i} : x = \sum_{i=1}^m c_i B_i^* x_i\right\} dx = \frac{I((f_i)_{i=1}^m)}{D} \end{aligned}$$

□

Lemma 9.4 can be extended to a more general setting compatible with the hypotheses of Theorem 9.1. By applying a similar approximation argument to the one used in the proof of Theorem 8.2, (26) holds when f_i and h_i are simple functions. Let $f_i, h_i \in L^1(\mathbb{R}^{n_i})$ be non-negative functions such that $\|f\|_1 = \|h\|_1$. There exists increasing sequences of simple functions $\phi_i^{(j)}$ and $\psi_i^{(j)}$ which converge to f_i and h_i respectively and satisfy

$\int_{\mathbb{R}^{n_i}} \phi_i^{(j)}(x)\,dx = \int_{\mathbb{R}^{n_i}} \psi_i^{(j)}(x)\,dx$ for all $j \in \mathbb{N}$. By Lemma 26,

$$I((f_i)_{i=1}^m) \geq I((\phi_i^{(n)})_{i=1}^m) \geq DJ((\psi_i^{(n)})_{i=1}^m) \to DJ((h_i)_{i=1}^m)$$

as $n \to \infty$. The limit is justified by the dominated convergence theorem, which may be applied because the integrand of $J((\psi_i^{(n)})_{i=1}^m)$ converges pointwise to and is dominated by the integrand of $J((h_i)_{i=1}^m)$.

Theorem 9.1 is a corollary of (26) for non-negative functions $f_i, h_i \in L^1(\mathbb{R}^{n_i})$. By taking the infimum over $\{f_i \in L^1(\mathbb{R}^{n_i}) : \|f_i\|_1 = 1, f_i \geq 0\}$ and the supremum over $\{h_i \in L^1(\mathbb{R}^{n_i}) : \|h_i\|_1 = 1, h_i \geq 0\}$ in (26), it follows that $E \geq DF$. Thus, by (24) and (25),

$$D^{\frac{1}{2}} = E_g \geq E \geq DF \geq DF_g = D^{\frac{1}{2}}$$

This implies that $E = D^{\frac{1}{2}}$ and $F = D^{-\frac{1}{2}}$, which completes the proof of Theorem 9.1.

This proof has two fundamental differences from the previous proofs of geometric and functional inequalities using the Brenier map. Firstly, this proof enables the Brascamp-Lieb inequality and its reverse form, the Barthe inequality, to be proved simultaneously, unlike the previous proofs which only prove one inequality. Secondly, the proof proceeds by computing the optimal constants E and F rather than by a direct computation leading to the inequalities in Theorem 9.1.

10 Discussion and Conclusion

Mass transport, and more specifically optimal transport, is the unifying feature of the proofs of the geometric and functional inequalities discussed. In the proof of the Brunn-Minkowski inequality by Hadwiger and Ohmann (1956), the cuts provide a method of transporting mass from one set to the other and localises the inequality to boxes within the sets. This conceptual idea motivates the application of the Brenier map, which localises the Brunn-Minkowski and isoperimetric inequalities to local determinant estimates in a more refined way (Alesker et al. 1999; Guillen and McCann 2010). The proof of the Prékopa-Leindler inequality due to McCann (1994) utilises the same localisation idea, albeit using the displacement interpolant between probability measures to provide an additional degree of freedom in the inequality. Conversely, Barthe (1998) applies the Brenier map differently in the simultaneous proof of the Brascamp-Lieb inequality and its reverse form, Barthe's inequality; his approach uses the Brenier map to transport mass between distributions which appear in each inequality, relating the two inequalities and enabling them to be proven together.

In conclusion, this paper has shown that the intertwined geometric and functional inequalities in Figure 1 are fundamentally related by their method of proof: applying transport maps such as the Brenier map to localise the inequality to a simpler one. As such, this paper extends on the survey by Gardner (2002), which primarily focused on how these inequalities are linked by their content.

11 Acknowledgements

I would like to thank my supervisor, A/Prof. Jiakun Liu, for his continual guidance and support. I would also like to thank AMSI and Jane Street for sponsoring this research project.

12 References

Alesker, S., Dar, S., and Milman, V. (1999). "A Remarkable Measure Preserving Diffeomorphism between two Convex Bodies in $\mathbb{R}^n$". In: *Geometriae Dedicata* 74.2, pp. 201–212. ISSN: 1572-9168. DOI: 10.1023/A:1005087216335.

Barthe, Franck (1998). "On a reverse form of the Brascamp-Lieb inequality". In: *Inventiones mathematicae* 134.2, pp. 335–361. ISSN: 1432-1297. DOI: 10.1007/s002220050267.

Evans, Lawrence C. (2001). *Partial Differential Equations and Monge–Kantorovich Mass Transfer*. Electronic Article. URL: https://math.berkeley.edu/~evans/Monge-Kantorovich.survey.pdf.

Figalli, A., Maggi, F., and Pratelli, A. (2009). "A refined Brunn–Minkowski inequality for convex sets". In: *Annales de l'Institut Henri Poincaré C, Analyse non linéaire* 26.6, pp. 2511–2519. ISSN: 0294-1449. DOI: 10.1016/J.ANIHPC.2009.07.004.

Gardner, Richard J. (2002). "The Brunn-Minkowski inequality". In: *Bulletin of the American Mathematical Society* 39.3, pp. 355–405. ISSN: 0273-0979.

Guillen, Nestor and McCann, Robert J. (2010). *Five lectures on optimal transportation: Geometry, regularity and applications*. Electronic Article. DOI: 10.48550/arXiv.1011.2911.

Hadwiger, H. and Ohmann, D. (1956). "Brunn-Minkowskischer Satz und Isoperimetrie". In: *Mathematische Zeitschrift* 66.1, pp. 1–8. ISSN: 1432-1823. DOI: 10.1007/BF01186589.

Hardy, G. H., Littlewood, J. E., and Pólya, G. (1934). *Inequalities*. London: Cambridge University Press.

McCann, Robert J. (1994). "A Convexity Theory for Interacting Gases and Equilibrium Crystals". Thesis.

13 Appendix

13.1 p-Means: Limiting Cases and Proof of Jensen's Inequality

The definition of the p-means (Definition 4.1) when $p = 0$ and $p = \pm\infty$ is consistent with the limiting behaviour of the definition when $p \neq 0$, where it is given by (2) (Hardy et al. 1934).

Consider the case where $p = \infty$. Without loss of generality, assume that $0 \leq a_1 \leq \cdots \leq a_n$. Then, for all $1 \leq i \leq n$, $\frac{a_i}{a_n} \leq 1$, so

$$\lambda_n^{\frac{1}{p}} a_n \leq \left(\sum_{i=1}^n \lambda_i a_i^p\right)^{\frac{1}{p}} = a_n \left(\sum_{i=1}^n \lambda_i \left(\frac{a_i}{a_n}\right)^p\right)^{\frac{1}{p}} \leq a_n \left(\sum_{i=1}^n \lambda_i\right)^{\frac{1}{p}} = a_n$$

By taking the limit $p \to \infty$ and applying the squeeze theorem, it follows that $\lim_{p\to\infty} M_p(\mathbf{a}, \boldsymbol{\lambda}) = a_n = \max_{1\leq i\leq n} a_i$, which is the definition of $M_\infty(\mathbf{a}, \boldsymbol{\lambda})$. The proof in the case where $p = -\infty$ is similar.

Consider the case where $p = 0$. Compute the logarithm of the p-mean.

$$\log\left(M_p(\mathbf{a}, \boldsymbol{\lambda})\right) = \frac{1}{p} \log\left(\sum_{i=1}^n \lambda_i e^{p\log a_i}\right)$$

Compute the limit as $p \to 0$ using L'Hôpital's Rule.

$$\lim_{p\to 0} \log\left(M_p(\mathbf{a}, \boldsymbol{\lambda})\right) = \lim_{p\to 0} \frac{\frac{d}{dp}\log\left(\sum_{i=1}^n \lambda_i e^{p\log a_i}\right)}{\frac{d}{dp}p} = \lim_{p\to 0} \frac{\sum_{i=1}^n \lambda_i e^{p\log a_i}\log a_i}{\sum_{i=1}^n \lambda_i e^{p\log a_i}} = \sum_{i=1}^n \lambda_i \log a_i$$

Thus, $\lim_{p\to 0} M_p(\mathbf{a}, \boldsymbol{\lambda}) = \exp(\sum_{i=1}^n \lambda_i \log a_i) = \prod_{i=1}^n a_i^{\lambda_i}$, which is the definition of $M_0(\mathbf{a}, \boldsymbol{\lambda})$.

Jensen's inequality for means (Theorem 4.1) can be proved by cases (Hardy et al. 1934).

Proof of Theorem 4.1. Consider the case where $0 < p < q$. Let $\beta = \frac{q}{p}$ and α be such that $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. Let $\mathbf{u} = \left(\lambda_i^{\frac{1}{\alpha}}\right)_{i=1}^n$ and $\mathbf{v} = \left(\lambda_i^{\frac{1}{\beta}} a_i^p\right)_{i=1}^n$. By the discrete Hölder inequality,

$$M_p(\mathbf{a}, \boldsymbol{\lambda}) = (\langle \mathbf{u}, \mathbf{v}\rangle)^{\frac{1}{p}} \leq (\|\mathbf{u}\|_\alpha \|\mathbf{v}\|_\beta)^{\frac{1}{p}} = \left(\sum_{i=1}^n \lambda_i\right)^{\frac{1}{p\alpha}} \left(\sum_{i=1}^n \lambda_i a_i^{p\beta}\right)^{\frac{1}{p\beta}} = \left(\sum_{i=1}^n \lambda_i a_i^q\right)^{\frac{1}{q}} = M_q(\mathbf{a}, \boldsymbol{\lambda})$$

Consider the case where $0 = p < q$. Assume that the weighted AM-GM inequality $M_0(\mathbf{a}, \boldsymbol{\lambda}) \leq M_1(\mathbf{a}, \boldsymbol{\lambda})$ holds. This is reasonable because, when $n = 2$, the weighted AM-GM inequality is equivalent to the concavity of the logarithm function, and when $n > 2$, the weighted AM-GM inequality can be proven by induction on n. Thus, by the weighted AM-GM inequality,

$$M_p(\mathbf{a}, \boldsymbol{\lambda}) = M_0(\mathbf{a}, \boldsymbol{\lambda}) = \left(\prod_{i=1}^n (a_i^q)^{\lambda_i}\right)^{\frac{1}{q}} \leq \left(\sum_{i=1}^n \lambda_i a_i^q\right)^{\frac{1}{q}} = M_q(\mathbf{a}, \boldsymbol{\lambda})$$

Consider the case where $p < q \leq 0$. Note the algebraic identity $M_{-r}(\mathbf{a}, \boldsymbol{\lambda}) = (M_r(\mathbf{a}^*, \boldsymbol{\lambda}))^{-1}$, where $\mathbf{a}^* = \left(\frac{1}{a_i}\right)_{i=1}^n$. Since Jensen's inequality has already been proven for $0 \leq -q < -p$, it follows that

$$M_p(\mathbf{a}, \boldsymbol{\lambda}) = \frac{1}{M_{-p}(\mathbf{a}^*, \boldsymbol{\lambda})} \leq \frac{1}{M_{-q}(\mathbf{a}^*, \boldsymbol{\lambda})} = M_q(\mathbf{a}, \boldsymbol{\lambda})$$

The remaining case where $p \leq 0 \leq q$ follows by combining the above cases. □

13.2 Approximation Arguments in the Proof of the General Brunn-Minkowski Inequality

When using the Lebesgue measure, any compact set can be approximated from the outside by finite unions of boxes and any non-empty, bounded, measurable set can be approximated from the inside by compact sets. These results form the basis of the approximation arguments used in the proof of the general Brunn-Minkowski inequality (Theorem 5.1) by Hadwiger-Ohmann cuts (Hadwiger and Ohmann 1956).

Lemma 13.1. Let X and Y be compact subsets of $\mathbb{R}^n$. Let $(X_i)_{i=1}^n$ and $(Y_i)_{i=1}^n$ be monotonically decreasing sequences of sets which can be decomposed into finite unions of boxes and satisfy $X_i \searrow X$ and $Y_i \searrow Y$ resepectively. Then, $X+Y$ is measurable, and $\lim_{i\to\infty} V(X_i+Y_i) = V(X+Y)$.

Proof. To show that $X+Y$ is measurable, consider the function $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ defined by $f(x,y) := x+y$, which is continuous. Then, $X+Y = f(X\times Y)$, which is the continuous image of a compact set, so it is compact and therefore measurable.

Since $X_i \searrow X$ and $Y_i \searrow Y$, it follows that $X = \bigcap_{i=1}^\infty X_i$ and $Y = \bigcap_{i=1}^\infty Y_i$. Note that for all $i \in \mathbb{N}$, X_i and Y_i are compact since finite unions of boxes are closed and bounded. Observe that $X+Y = \bigcap_{i=1}^\infty (X_i+Y_i)$. To see this, let $z \in X+Y$, $z = x+y$ for some $x \in X$ and $y \in Y$. Then, $x \in X_i$ and $y \in Y_i$ for all $i \geq 1$, so $z = x+y \in \bigcap_{i=1}^\infty (X_i+Y_i)$. Now, let $z \in \bigcap_{i=1}^\infty (X_i+Y_i)$. Then, for each $i \geq 1$, there exists some $x_i \in X_i$ and $y_i \in Y_i$ such that $z = x_i + y_i$. Since $(X_i)_{i=1}^n$ and $(Y_i)_{i=1}^n$ are monotonically decreasing, $(x_i)_{i=1}^\infty \subset X_1$ and $(y_i)_{i=1}^\infty \subset Y_1$, so, by compactness, there exists convergent subsequences $(x_{i_j})_{j=1}^\infty$ and $(y_{i_j})_{i=1}^\infty$. Let $x := \lim_{j\to\infty} x_{i_j}$ and $y := \lim_{j\to\infty} y_{i_j}$ and fix $n \in \mathbb{N}$. There exists some j such that $x_{i_k} \in X_n$ $(y_{i_k} \in Y_n)$ for all $k \geq j$, so $x \in X_n$ $(y \in Y_n)$ by compactness. Since n is arbitrary, it follows that $x \in X$ and $y \in Y$, so by taking limits, $z = \lim_{j\to\infty}(x_{i_j} + y_{i_j}) = x+y \in X+Y$.

The volume limit follows naturally.

$$\lim_{i\to\infty} V(X_i+Y_i) = \lim_{i\to\infty} V\left(\bigcap_{j=1}^{i}(X_j+Y_j)\right) = V\left(\bigcap_{j=1}^{\infty}(X_j+Y_j)\right) = V(X+Y)$$

□

Lemma 13.2. Let X and Y be measurable subsets of $\mathbb{R}^n$ such that $X+Y$ is measurable. Let $(X_i)_{i=1}^n$ and $(Y_i)_{i=1}^n$ be monotonically increasing sequences of compact sets which satisfy $X_i \nearrow X$ and $Y_i \nearrow Y$ respectively. Then, $\lim_{i\to\infty} V(X_i+Y_i) = V(X+Y)$.

Proof. Since $X_i \nearrow X$ and $Y_i \nearrow Y$, it follows that $X = \bigcup_{i=1}^\infty X_i$ and $Y = \bigcup_{i=1}^\infty Y_i$. By the distributive property of the Minkowski sum over unions,

$$X+Y = \left(\bigcup_{i=1}^{\infty} X_i\right) + \left(\bigcup_{i=1}^{\infty} Y_i\right) = \bigcup_{i=1}^{\infty}\bigcup_{j=1}^{\infty}(X_i+Y_j) = \bigcup_{i=1}^{\infty}(X_i+Y_i)$$

where the diagonal terms in the double union can be omitted because $X_i + Y_j \subset X_{\max\{i,j\}} + Y_{\max\{i,j\}}$.

The volume limit follows naturally.

$$\lim_{i\to\infty} V(X_i+Y_i) = \lim_{i\to\infty} V\left(\bigcup_{j=1}^{i}(X_j+Y_j)\right) = V\left(\bigcup_{j=1}^{\infty}(X_j+Y_j)\right) = V(X+Y)$$

□

13.3 Technical Lemmas for Determinants

Lemma 13.3. The determinant operator $\det : A \mapsto \det(A)$ is $\frac{1}{n}$-concave on the space of symmetric and positive semidefinite matrices. That is, if $A, B \in \mathbb{R}^{n\times n}$ are symmetric and positive semidefinite,

$$\det(A+B)^{\frac{1}{n}} \geq \det(A)^{\frac{1}{n}} + \det(B)^{\frac{1}{n}} \tag{30}$$

Equality holds in (30) when $A = cB$ for some $c \geq 0$

Proof. If $\det(A) = \det(B) = 0$, (30) follows automatically since $A+B$ is symmetric and positive semidefinite.

Assume that $\det(B) > 0$. We use the method presented in Figalli et al. (2009). By the multiplicative property of the determinant and the spectral theorem for real symmetric matrices, it can be assumed without loss of generality that A is diagonal, and $B = I$. Let $\lambda_1, \ldots, \lambda_n$ be the diagonal entries of A. Thus, it is sufficient to show that

$$\prod_{i=1}^{n}(\lambda_i+1) = \det(A+B) \geq \left(\det(A)^{\frac{1}{n}} + \det(B)^{\frac{1}{n}}\right)^n = \left(\left(\prod_{i=1}^{n}\lambda_i\right)^{\frac{1}{n}} + 1\right)^n$$

Expand $\prod_{i=1}^{n}(\lambda_i+1)$ as

$$\prod_{i=1}^{n}(\lambda_i+1) = 1 + \sum_{m=1}^{n}\sum_{(i_1,\ldots,i_m)\in I}\prod_{j=1}^{m}\lambda_{i_j}$$

where $I = \{(i_1,\ldots,i_m) : 1 \leq i_1 < \cdots < i_m \leq n\}$ is the set of multi-indices, which contains $\binom{n}{m}$ elements. By the AM-GM inequality,

$$\sum_{(i_1,\ldots,i_m)\in I}\prod_{j=1}^{m}\lambda_{i_j} \geq \binom{n}{m}\prod_{(i_1,\ldots,i_m)\in I}\left(\prod_{j=1}^{m}\lambda_{i_j}\right)^{\frac{1}{\binom{n}{m}}} = \binom{n}{m}\prod_{(i_1,\ldots,i_m)\in I}\prod_{j=1}^{m}\lambda_{i_j}^{\frac{1}{\binom{n}{m}}}$$

For each $1 \leq k \leq n$, there are $\binom{n-1}{m-1}$ elements in I which contain k. Thus,

$$\prod_{(i_1,\ldots,i_m)\in I}\prod_{j=1}^{m}\lambda_{i_j}^{1/\binom{n}{m}} = \prod_{k=1}^{m}\lambda_k^{\binom{n-1}{m-1}/\binom{n}{m}} = \left(\prod_{k=1}^{m}\lambda_k\right)^{\frac{m}{n}}$$

The above results imply

$$\prod_{i=1}^{n}(\lambda_i+1) \geq \sum_{i=1}^{n}\binom{n}{m}\left(\prod_{k=1}^{m}\lambda_k\right)^{\frac{m}{n}} = \left(\left(\prod_{k=1}^{m}\lambda_k\right)^{\frac{1}{n}} + 1\right)^n$$

By the equality conditions of the AM-GM inequality, equality holds in (30) when $\lambda_1 = \cdots = \lambda_n := c$, which requires that $A = cI = cB$. □

Lemma 13.4. Let $A \in \mathbb{R}^{n\times n}$ be symmetric and positive semidefinite. Then,

$$\frac{1}{n}\operatorname{tr}(A) \geq \det(A)^{\frac{1}{n}}$$

Equality holds when $A = cI$ for some $c \geq 0$.

Proof. Let the eigenvalues of A be $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$. Then, by the AM-GM inequality,

$$\frac{1}{n}\operatorname{tr}(A) = \frac{1}{n}\sum_{i=1}^{n}\lambda_i \leq \prod_{i=1}^{n}\lambda_i = \det(A)$$

By the equality conditions of the AM-GM inequality, equality holds when $\lambda_1 = \cdots = \lambda_n := c$, which requires that $A = cI$. □

13.4 Technical Lemmas in the Proof of the Borell-Brascamp-Lieb Inequality

Proof of Lemma 8.4. Since Remark 8.3 implies that $D^2\phi$ exists, the Jacobian of s_t is $Ds_t(x) = (1-t)I + tD^2\phi(x)$ for all $x \in \Omega_{f+}$. Moreover, for any $y \in \mathbb{R}^n \setminus \{0\}$, Ds_t satisfies

$$\langle Ds_t(x)y, y\rangle = (1-\lambda)\|y\|^2 + \lambda\left\langle D^2\phi(x)y, y\right\rangle \geq (1-\lambda)\|y\|^2 > 0$$

since $D^2\phi(x)$ is positive semidefinite as it is the Hessian of a convex potential. This proves that $Ds_t(x)$ is positive definite, so $\det Ds_t(x) > 0$. Therefore, by the inverse function theorem, s_t is a local C^1-diffeomorphism.

Let $x, y \in \mathbb{R}^n$, $x \neq y$. Since ϕ is convex, $\phi(x) \geq \phi(y) + \langle D\phi(y), x-y\rangle$ and $\phi(y) \geq \phi(x) + \langle D\phi(x), y-x\rangle$, which implies that

$$\langle x-y, D\phi(x) - D\phi(y)\rangle = \langle x-y, D\phi(x)\rangle - \langle x-y, D\phi(y)\rangle \geq (\phi(x)-\phi(y)) - (\phi(x)-\phi(y)) = 0$$

Therefore,

$$\langle x-y, s_t(x) - s_t(y)\rangle = (1-\lambda)\|x-y\|^2 + \lambda\langle x-y, D\phi(x) - D\phi(y)\rangle \geq (1-\lambda)\|x-y\|^2 > 0$$

so $s_t(x) \neq s_t(y)$, which implies by contraposition that s_t is injective. It follows that s_t is a C^1-diffeomorphism onto its image, so it has a differentiable inverse $s_t^{-1} : s_t(\Omega_{f+}) \to \Omega_{f+}$.

By (17) and applying the substitution $x = s_t(y)$

$$\int_{s_t(\Omega_{f+})} \xi(x)\,d\rho_t(x) = \int_{\Omega_{f+}} \xi(s_t(y))f(y)\,dy = \int_{s_t(\Omega_{f+})} \xi(x)f(s_t^{-1}(x))\det Ds_t^{-1}(x)\,dx \tag{31}$$

for every continuous function ξ. By the inverse function theorem, $Ds_t^{-1}(x) = (Ds_t(s_t^{-1}(x)))^{-1}$. Thus, the probability measure ρ_t admits a density function $\rho_t(x; f^+, f^-)$ which satisfies

$$\rho_t(x; f^+, f^-) = f^+(s_t^{-1}(x))\det(Ds_t(s_t^{-1}(x)))^{-1}\,dx \tag{32}$$

for almost every x, with respect to the Lebesgue measure. □

Lemma 13.5. Let $f, g \in L^1(\mathbb{R}^n)$ be such that $\|f\|_1 \neq 0$ and $\|g\|_1 \neq 0$. Define the normalised functions $\tilde{f} = \frac{f}{\|f\|_1}$ and $\tilde{g} = \frac{g}{\|g\|_1}$. Let $S_k : L^1(\mathbb{R}^n) \to L^1(\mathbb{R}^n)$ be the volume-preserving scaling operator defined by $(S_k f)(x) := k^{-n} f\left(\frac{x}{k}\right)$. Fix $p > -\frac{1}{n}$ and define the constants $\gamma := \frac{p}{np+1}$ and $C := ((1-\lambda)\|f\|_1^\gamma + \lambda\|g\|_1^\gamma)^{\frac{1}{\gamma}}$ (the right hand side of (16)). Then,

$$h_p(x; f, g, \lambda) = C h_p\left(x; S_{C_1^\gamma}\tilde{f}, S_{C_2^\gamma}\tilde{g}; C_2^{-\gamma}\lambda\right) \tag{33}$$

where h_p is defined by (19), $C_1 := \frac{C}{\|f\|_1}$ and $C_2 := \frac{C}{\|g\|_1}$.

Proof. By definition,

$$\frac{1-\lambda}{C_1^\gamma} + \frac{\lambda}{C_2^\gamma} = \frac{(1-\lambda)\|f\|_1^\gamma + \lambda\|g\|_1^\gamma}{C^\gamma} = 1 \tag{34}$$

Let $z_1, z_2 \in \mathbb{R}^n$. If $p \in \left(-\frac{1}{2}, \infty\right] \setminus \{0, \infty\}$, it follows from the definitions of the p-mean and γ and (34) that

$$\begin{aligned} M_p(f(z_1), g(z_2); \lambda) := ((1-\lambda)f(z_1)^p + \lambda g(z_2)^p)^{\frac{1}{p}} &= \left(\frac{1-\lambda}{C_1^{p(\gamma n-1)+\gamma}} f(z_1)^p + \frac{\lambda}{C_2^{p(\gamma n-1)+\gamma}} g(z_2)^p\right)^{\frac{1}{p}} \\ &= \left((1 - C_2^{-\gamma}\lambda)\left(C_1^{1-\gamma n} f(z_1)\right)^p + C_2^{-\gamma}\lambda\left(C_2^{1-\gamma n} g(z_2)\right)^p\right)^{\frac{1}{p}} \\ &= M_p\left(C_1^{1-\gamma n} f(z_1), C_2^{1-\gamma n} g(z_2); C_2^{-\gamma}\lambda\right) \end{aligned} \tag{35}$$

(35) also holds if $p = 0$ or $p = \infty$ by considering the limiting behaviour of the p-mean in each case.

By computing $h_p(x; f, g, \lambda)$ using (35)

$$\begin{aligned} h_p(x; f, g, \lambda) &= \sup\{M_p(f(z_1), g(z_2); \lambda) : (1-\lambda)z_1 + \lambda z_2 = x\} \\ &= \sup\left\{M_p\left(C_1^{1-\gamma n} f(z_1), C_2^{1-\gamma n} g(z_2); C_2^{-\gamma}\lambda\right) : (1-\lambda)z_1 + \lambda z_2 = x\right\} \\ &= C\sup\left\{M_p\left(C_1^{-\gamma n}\tilde{f}(z_1), C_2^{-\gamma n}\tilde{g}(z_2); C_2^{-\gamma}\lambda\right) : (1-\lambda)z_1 + \lambda z_2 = x\right\} \end{aligned}$$

By applying the substitutions $z_1 = C_1^{-\gamma} y_1$ and $z_2 = C_1^{-\gamma} y_2$ and using (34),

$$\begin{aligned} h_p(x; f, g, \lambda) &= C\sup\left\{M_p\left(C_1^{-\gamma n}\tilde{f}(C_1^{-\gamma} y_1), C_2^{-\gamma n}\tilde{g}(C_2^{-\gamma} y_2); C_2^{-\gamma}\lambda\right) : C_1^{-\gamma}(1-\lambda)y_1 + C_2^{-\gamma}\lambda y_2 = x\right\} \\ &= C\sup\left\{M_p\left((S_{C_1^\gamma}\tilde{f})(y_1), (S_{C_2^\gamma}\tilde{g})(y_2); C_2^{-\gamma}\lambda\right) : (1 - C_2^{-\gamma}\lambda)y_1 + C_2^{-\gamma}\lambda y_2 = x\right\} \\ &= C h_p\left(x; S_{C_1^\gamma}\tilde{f}, S_{C_2^\gamma}\tilde{g}, C_2^{-\gamma}\lambda\right) \end{aligned}$$

□

13.5 Technical Lemmas in the Proof of the Brascamp-Lieb Inequality and Barthe's Inequality

We first derive a formula for the integral of a centred Gaussian. Let $A \in \mathbb{R}^{n\times n}$ be a symmetric and positive definite matrix. By a change of coordinates if necessary, it can be assumed that A is diagonal with respect to the standard basis of $\mathbb{R}^n$. Let the diagonal entries of A be $\lambda_1, \ldots, \lambda_n > 0$. Then,

$$\begin{aligned}\int_{\mathbb{R}^n} G_A(x)\,dx := \int_{\mathbb{R}^n} \exp(-\langle Ax, x\rangle)\,dx &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left(-\sum_{i=1}^n \lambda_i x_i^2\right) dx_1 \ldots dx_n \\ &= \prod_{i=1}^n \int_{-\infty}^{\infty} e^{-\lambda_i x_i^2}\,dx_i \\ &= \prod_{i=1}^n \sqrt{\frac{\pi}{\lambda_i}} \\ &= \sqrt{\frac{\pi^n}{\det A}} \end{aligned} \tag{36}$$

This formula will be used to compute F_g and E_g.

Proof of Lemma 9.2. Compute the numerator inside the supremum by applying (36).

$$\begin{aligned} J((G_{A_i})_{i=1}^m) := \int_{\mathbb{R}^n} \prod_{i=1}^m G_{A_i}(B_i x)^{c_i}\,dx &= \int_{\mathbb{R}^n} \prod_{i=1}^m e^{-c_i\langle AB_i x, B_i x\rangle}\,dx \\ &= \int_{\mathbb{R}^n} \exp\left(-\left\langle \sum_{i=1}^m c_i B_i^* AB_i x, x\right\rangle\right) dx \\ &= \sqrt{\frac{\pi^n}{\det\left(\sum_{i=1}^m c_i B_i^* AB_i\right)}} \end{aligned} \tag{37}$$

Compute the denominator inside the supremum by applying (36) and using the assumption $\sum_{i=1}^m c_i n_i = n$.

$$\prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} G_{A_i}(x)\,dx\right)^{c_i} = \prod_{i=1}^m \left(\frac{\pi^{n_i}}{\det A_i}\right)^{\frac{c_i}{2}} = \sqrt{\frac{\pi^n}{\prod_{i=1}^m (\det A_i)^{c_i}}} \tag{38}$$

(24) follows by combining (37) and (38) and applying the definitions of F_g and D. □

Proof of Lemma 9.3. Let $x = \sum_{i=1}^m c_i B_i^* x_i$, where $x_i \in \mathbb{R}^{n_i}$. The matrices A_i for $1 \le i \le m$ have invertible, self-adjoint square roots $A_i^{\frac{1}{2}}$ because they are symmetric and positive definite. Thus,

$$|\langle x, y\rangle|^2 = \left|\left\langle \sum_{i=1}^m c_i B_i^* x_i, y\right\rangle\right|^2 = \left|\sum_{i=1}^m \langle c_i x_i, B_i y\rangle\right|^2 = \left|\sum_{i=1}^m \left\langle \sqrt{c_i} A_i^{-\frac{1}{2}} x_i, \sqrt{c_i} A_i^{\frac{1}{2}} B_i y\right\rangle\right|^2$$

The sum over the inner products on $\mathbb{R}^{n_i}$ can be interpreted as an inner product on $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$.

$$|\langle x, y\rangle|^2 = \left|\sum_{i=1}^m \left\langle \sqrt{c_i} A_i^{-\frac{1}{2}} x_i, \sqrt{c_i} A_i^{\frac{1}{2}} B_i y\right\rangle\right|^2 = \left|\left\langle \left(\sqrt{c_i} A_i^{-\frac{1}{2}} x_i\right)_{i=1}^m, \left(\sqrt{c_i} A_i^{\frac{1}{2}} B_i y\right)_{i=1}^m \right\rangle\right|^2 \tag{39}$$

By the Cauchy-Schwarz inequality on $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$,

$$\begin{aligned} |\langle x, y\rangle|^2 &\le \left\|\left(\sqrt{c_i} A_i^{-\frac{1}{2}} x_i\right)_{i=1}^m\right\|^2 \left\|\left(\sqrt{c_i} A_i^{\frac{1}{2}} B_i y\right)_{i=1}^m\right\|^2 \\ &= \left(\sum_{i=1}^m \left\langle \sqrt{c_i} A_i^{-\frac{1}{2}} x_i, \sqrt{c_i} A_i^{-\frac{1}{2}} x_i\right\rangle\right)\left(\sum_{i=1}^m \left\langle \sqrt{c_i} A_i^{\frac{1}{2}} B_i y, \sqrt{c_i} A_i^{\frac{1}{2}} B_i y\right\rangle\right) \\ &= \left(\sum_{i=1}^m c_i \left\langle A_i^{-1} x_i, x_i\right\rangle\right)\left\langle \sum_{i=1}^m c_i B_i^* A_i B_i y, y\right\rangle \end{aligned}$$

Define $K((A_i)_{i=1}^m) := \inf\left\{\sum_{i=1}^m c_i \left\langle A_i^{-1}x_i, x_i\right\rangle : x = \sum_{i=1}^m c_i B_i^* x_i\right\}$ and note that it satisfies

$$|\langle x,y\rangle|^2 \leq K((A_i)_{i=1}^m)\left\langle \sum_{i=1}^m c_i B_i^* A_i B_i y, y\right\rangle \leq \left(\sum_{i=1}^m c_i \left\langle A_i^{-1}x_i, x_i\right\rangle\right)\left\langle \sum_{i=1}^m c_i B_i^* A_i B_i y, y\right\rangle \tag{40}$$

Set $y = \left(\sum_{i=1}^m c_i B_i^* A_i B_i\right)^{-1} x$ and $x_i = A_i B_i y$. Then, $\sqrt{c_i}A_i^{-\frac{1}{2}}x_i = \sqrt{c_i}A_i^{\frac{1}{2}}B_i y$ for all $1 \leq i \leq m$, so the equality condition of the Cauchy-Schwarz inequality on $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$ is satisfied. Thus, (40) achieves equality, so $\langle x,y\rangle$ solves the quadratic equation $\langle x,y\rangle^2 - K((A_i)_{i=1}^m)\langle x,y\rangle = 0$. This equation has two solutions: $\langle x,y\rangle = 0$ and $\langle x,y\rangle = K((A_i)_{i=1}^m)$. By (39), $\langle x,y\rangle = \left\|\left(\sqrt{c_i}A_i^{\frac{1}{2}}B_i y\right)_{i=1}^m\right\|$, a norm on $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$, so $\langle x,y\rangle = 0$ if and only if $x = y = 0$ by the positive definiteness of the norm. However, if $x = 0$, $K((A_i)_{i=1}^m) = 0 = \langle x,y\rangle$, so for all $x \in \mathbb{R}^n$,

$$K((A_i)_{i=1}^m) = \langle x,y\rangle = \left\langle \left(\sum_{i=1}^m c_i B_i^* A_i B_i\right)^{-1} x, x\right\rangle \tag{41}$$

Now, compute $I((G_{A_i^{-1}})_{i=1}^m)$.

$$\begin{aligned} I((G_{A_i^{-1}})_{i=1}^m) &= \overline{\int_{\mathbb{R}^n}} \sup\left\{\prod_{i=1}^m G_{A_i^{-1}}(x_i)^{c_i} : x = \sum_{i=1}^m c_i B_i^* x_i,\ x_i \in \mathbb{R}^{n_i}\right\} dx \\ &= \overline{\int_{\mathbb{R}^n}} \sup\left\{\prod_{i=1}^m \exp(-c_i \left\langle A_i^{-1}x_i, x_i\right\rangle) : x = \sum_{i=1}^m c_i B_i^* x_i,\ x_i \in \mathbb{R}^{n_i}\right\} dx \\ &= \overline{\int_{\mathbb{R}^n}} \exp\left(-\inf\left\{\sum_{i=1}^m c_i \left\langle A_i^{-1}x_i, x_i\right\rangle : x = \sum_{i=1}^m c_i B_i^* x_i,\ x_i \in \mathbb{R}^{n_i}\right\}\right) dx \end{aligned}$$

By the definition of $K((A_i)_{i=1}^m)$ and by applying (41) and (36),

$$I((G_{A_i^{-1}})_{i=1}^m) = \overline{\int_{\mathbb{R}^n}} \exp\left(-\left\langle \left(\sum_{i=1}^m c_i B_i^* A_i B_i\right)^{-1} x, x\right\rangle\right) dx = \sqrt{\pi^n \det\left(\sum_{i=1}^m c_i B_i^* A_i B_i\right)} \tag{42}$$

where the upper Lebesgue integral is computed as the Lebesgue integral since the integrand is integrable. By taking the infimum in (25) over A_i^{-1} instead of A_i, which is possible since A_i is invertible, it follows that

$$E_g = \inf\left\{\frac{I((G_{A_i^{-1}})_{i=1}^m)}{\prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} G_{A_i^{-1}}(x)\,dx\right)^{c_i}} : A_i \in \mathbb{R}^{n\times n} \text{ is symmetric and positive definite}\right\}$$

(25) follows by combining (38) and (42) and applying the definitions of E_g and D. □