Convex measures and associated geometric and functional inequalities

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

James Melbourne

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

SERGEY BOBKOV

July, 2015

Acknowledgements

This thesis was partially supported by NSF grant DMS-1106530 and the University of Minnesota Graduate School's Thesis Research Travel Grant.

Abstract

Convex measures represent a class of measures that satisfy a variant of the classical Brunn-Minkowski Inequality. Background on the associated functional and geometric inequalities is given, and the elementary theory of such measures is explored. A generalization of the Lovasz and Simonovits localization technique is developed, and some applications to large deviations are explained. In a more geometric direction, a modified Brunn-Minkowski Inequality is explored on some discrete spaces. The significance of such a notion is in its potential to serve as a definition for a lower Ricci curvature bound in non-smooth spaces.

Contents

Chapter 1

Introduction

The general theory of so-called convex measures on locally convex spaces E was created in the mid 1970's by C. Borell. One of the motivations of Borell was the desire to study a number of general properties of Gaussian measures (and Gaussian processes) by means of dimension-free geometric inequalities. He also noticed that the distributions of some other random processes related to the Brownian motion possess certain convexity properties. Another source of motivation comes from Convex Geometry: distributions of linear functionals over convex bodies turn out to be convex in some sense; more precisely, they are log-concave regardless of the dimension of the body.

Let us note that any Gaussian measure is log-concave, as well. Previously, the class of log-concave measures was considered by several authors, and one should mention the works by Prékopa and Leindler in early 1970's. With the help of Brunn-Minkowski-type geometric inequalities, Borell introduced more general classes of α-concave measures with parameter of convexity $\alpha \in \mathbb{R}$. Many properties of such measures depend on the parameter α, only, and do not depend on the dimension of the space E. This allows one to extend various finite dimensional theorems about convex measures to spaces of an infinite dimnension, and this way one may study a number of global properties of distributions of random processes (often treated as probability measures on classical functional spaces). Borell has studied most general properties of convex probability measures, such as the 0-1 law, integrability of norms, convexity preserving under convolutions. He also gave a full characterization of such measures on $\mathbb{R}^n$ in terms of their densities.

A further deep investigation of convex measures (especially in the log-concave case) was started by Lovász and Simonovits in the mid 1990's. They introduced and developed a new approach, known nowadays as localization, which allows one to reduce various multidimensional integral relations to dimension 1 (they attribute main ideas for this approach to Payne and Weinberger who considered Poincaré-type inequalities on convex bodies). Localization thechnique turned out to be a rather powerful tool in the whole theory of convex measures. In particular, it allows to attack different problems including obtaining sharp dilation and Khinchine-type estimates, as well as isoperimetric and Sobolev-type inequalities over convex measures. For example, as was shown by Bobkov, with this approach one can easily recover the Gaussian isoperimetric inequality of Borell-Sudakov-Tsirel'son, and to obtain its generalizations including the Bakry-Ledoux isoperimetric inequality. There are also other, alternative lines of investigation of convex measures; they are usually based on the application of the Prékopa-Leindler theorem and its dimensional extension due to Brascamp and Lieb, or are based on the application of suitable transference plans (the transport approach).

In Chapters 2 and 3 of this thesis we remind basic notions in the theory of convex measures and describe main tools of investigation. One of our main further purposes (Chapter 4) is to extend the localization technique to spaces of an infinite dimension. In particular, the so-called bisection method is developed for abstract Frechet spaces. There have been also obtained an infinite dimensional extension of Fradelizi-Guedon's theorem on extreme convex measures supported on convex sets of arbitrary complete locally convex spaces. Although the results of this type are known for Euclidean spaces $\mathbb{R}^n$, the desired extensions are not immediate and are not simple, as might seem at the first sight. Then we describe several applications to dilation-type inequalities on infinite dimensional spaces, which generalize corresponding inequalities due to Nazarov, Sodin and Volberg (2002), Bobkov, Nazarov (2007), and Fradelizi (2008).

These results were reported on at several conferences and have been published in "Doklady of Russian Academy of Sciences" [3] with a forth coming publication in the journal "Probability Surveys" [4] both with S. Bobkov. The work has already found several applications in the study of distributions of polynomials over convex measures [1, 2] (Bogachev's school in Moscow).

Chapter 5 is devoted to Brunn-Minkowski-type inequalities on graphs and their connections with a generalized notion of the Ricci curvature tensor. These ideas represent discrete analogs of the synthetic Ricci theory put forth by Villani, Strum, Otto et al, background can be found in [39] while the discrete Brunn Minkowski was initiated in [34] and relationships to functional inequalities established via the related displacement convexity of entropy can be found in [25].

This section is part of an ongoing project with N. Gozlan, W. Perkins, C. Roberto, P. Tetali, and P.M. Samson to develop and understand, in the discrete setting, the role of geometry in concentration of measure phenomena. With the authors above, a survey/manuscript on discrete curvature is currently in preparation [23], some of the contributions contained in this thesis will be part of the in preparation article [24].

Chapter 2

Brunn-Minkowski Inequality

For (X, μ) a metric measure space, we will be interested in inequalities that relate the size of a spacial average of two sets, to an average of the sizes of the two sets. This notion is pervasive and fundamental, manifesting in numerous fields and contexts. Before introducing further generality, and definitions, let us start with an example.

2.1 Brunn-Minkowski Inequality in $\mathbb{R}^n$

The Brunn-Minkowski inequality takes several equivalent forms in $\mathbb{R}^n$. It can be used to derive isoperimetric and concentration inequalities. We will use the following notation for the Minkowski sum of two sets and the set dialation, for $A, B \subset \mathbb{R}^n$, $t \in \mathbb{R}$.

$$A + B = \{a + b : a \in A, b \in B\} \tag{2.1}$$

$$tA = \{ta : a \in A\}. \tag{2.2}$$

The above definitions make sense on any real vector space, and are enough to introduce the aforementioned "set theoretic average" of two sets in a vector space context. We can now state **the** Brunn-Minkowski inequality.

Theorem 2.1.1 *For two Borel measurable sets $A, B \subset \mathbb{R}^n$,*

$$|(1-t)A + tB| \geq \left((1-t)|A|^{\frac{1}{n}} + t|B|^{\frac{1}{n}}\right)^n.$$

There are several equivalent formulations.

Claim:

If any of the following hold for all A, B Borel,

$$|(1-t)A + tB| \geq \left((1-t)|A|^{\frac{1}{n}} + t|B|^{\frac{1}{n}}\right)^n \tag{2.3}$$

$$|(1-t)A + tB| \geq |A|^{1-t}|B|^t \tag{2.4}$$

$$|(1-t)A + tB| \geq \min\{|A|, |B|\} \tag{2.5}$$

$$|A + B| \geq \left(|A|^{\frac{1}{n}} + |B|^{\frac{1}{n}}\right)^n \tag{2.6}$$

the all of the above hold.

Proof:

That the first statement implies the second, which implies the third follows from an inequality sometimes attributed to Markov. It is as follows, for a probability space with expectation denoted $\mathbb{E}$, and random variable X, the function $f(p) = (\mathbb{E}X^p)^{\frac{1}{p}}$ is non-decreasing. The theorem's proof is an application of Jensen's inequality, to see its relation to the above let X take value $|A|$ with probability $1-t$ and $|B|$ with probability t.

To show that the third statement implies the fourth, let $t \in (0, 1)$

$$\begin{aligned}
|A + B| =& |(1-t)(A/(1-t)) + t(B/t)| \\
\geq& \min\{|A/(1-t)|, |B/t|\} \\
=& \min\{\frac{|A|}{(1-t)^n}, \frac{|B|}{t^n}\}.
\end{aligned}$$

Solving $\frac{|A|}{(1-t)^n} = \frac{|B|}{t^n}$ for t we get $t = |B|^{1/n}/(|A|^{1/n} + |B|^{1/n})$. Inserting this into the above, we have 4.

Assuming the forth statement,

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

and the desired equivalence is proven.

2.1.1 Prékopa-Leindler

The Brunn-Minkowski inequality has a functional variant see [18] which we will prove equivalent in $\mathbb{R}^n$.

Theorem 2.1.2 *Given* $f, g, h : \mathbb{R}^n \to [0, \infty)$ *and a* $t \in [0, 1]$ *such that for any* $x, y \in \mathbb{R}^n$, $f((1-t)x + ty)) \geq g^{1-t}(x)h^t(y)$ *holds, then*

$$\int_{\mathbb{R}^n} f \geq \left(\int_{\mathbb{R}^n} g\right)^{1-t} \left(\int_{\mathbb{R}^n} h\right)^t. \tag{2.7}$$

We will prove Brunn-Minkowski directly in one dimension, and then use this result and an induction argument to prove Prékopa-Leindler from which the n dimensional statement of Brunn-Minkowski will also follow.

Proof We wish to first show that for nonempty A, B Borel $\subset \mathbb{R}$, $|A + B| \geq |A| + |B|$. First note that by the the regularity of the Lebesgue measure, it suffices to assume A and B are compact. By translation invariance we may also assume that $\min\{b \in B\} = 0 = \max\{a \in A\}$. With these assumptions $A \cup B \subset A + B$ and $A \cap B = \{0\}$ has measure zero, so that

$$|A + B| \geq |A \cup B| = |A| + |B|.$$

Now assume f, g, h are functions on $\mathbb{R}$ that satisfy the hypothesis of Prékopa-Leindler inequality. By homogeneity we may assume that $esssupf = esssupg = 1$ (this normalization is only to ensure that we avoid the application of Brunn-Minkowski Inequality to empty sets.

$$\int_{\mathbb{R}} f = \int_0^\infty |\{x : f(x) > y\}| dy \tag{2.8}$$

$$= \int_0^1 |\{x : f(x) > y\}| dy \tag{2.9}$$

$$\geq \int_0^1 |(1-t)\{g > y\} + t\{h > y\}| dy \tag{2.10}$$

$$\geq (1-t) \int_0^1 |\{g > y\}| dy + t \int_0^1 |\{h > y\}| dy \tag{2.11}$$

$$\geq \left(\int g\right)^{1-t} \left(\int h\right)^t. \tag{2.12}$$

The first equality follows from Fubini-Tonelli, the first inequality holds via set theoretic inclusion, the second is Brunn Minkowski, the last is the same Fubini-Tonelli combined with Arithmetic/Geometric mean inequality.

Tensorization of Prékopa-Leindler (PL)

Now by induction, suppose that PL holds on $\mathbb{R}^n$ and that f, g, h are functions on $\mathbb{R}^{n+1}$ that satisfy the hypothesis of PL. Now define F, G, H on $\mathbb{R}$ by

$$\begin{aligned} F(s) &= \int_{\mathbb{R}^n} f(x,s)dx, \\ G(s) &= \int_{\mathbb{R}^n} g(x,s)dx, \\ H(s) &= \int_{\mathbb{R}^n} h(x,s)dx. \end{aligned}$$

For a fixed $s \in \mathbb{R}$ define $f_s(x) = f(x,s),\ g_s(x) = g(x,s), h_s(x) = h(x,s)$. Writing in coordinates $\mathbb{R}^n \times \mathbb{R}$ the fact that f, g, h satisfy the (PL) hypothesis is the following

$$f((1-t)(x,s) + t(y,r)) \geq g^{1-t}(x,s)h^t(y,r)$$

. With our newly defined functions we can write this as

$$f_{(1-t)s+tr}((1-t)x + ty) \geq g_s^{1-t}(x)h_r^t(y).$$

This is exactly the statement that $f_{(1-t)s+tr}, g_s, h_r$ satisfy the hypothesis of PL on $\mathbb{R}^n$. Via the induction

$$\int_{\mathbb{R}^n} f_{(1-t)s+tr}(z)dz \geq \left(\int_{\mathbb{R}^n} g_s(z)dz\right)^{1-t} \left(\int_{\mathbb{R}^n} h_r(z)dz\right)^t$$

Making note of our definitions this is the fact that

$$F((1-t)s + tr) \geq G^{1-t}(s)H^t(r).$$

Applying PL on $\mathbb{R}$,

$$\begin{gathered} \int_{\mathbb{R}} F(t)dt \geq \left(\int_{\mathbb{R}} G(t)dt\right)^{1-t} \left(\int_{\mathbb{R}} H(t)dt\right)^t \\ \int_{\mathbb{R}}\int_{\mathbb{R}^n} f(x,t)dxdt \geq \left(\int_{\mathbb{R}}\int_{\mathbb{R}^n} g(x,t)dxdt\right)^{1-t} \left(\int_{\mathbb{R}}\int_{\mathbb{R}^n} h(x,t)dxdt\right)^t \\ \int_{\mathbb{R}^{n+1}} f \geq \left(\int_{\mathbb{R}^{n+1}} g\right)^{1-t} \left(\int_{\mathbb{R}^{n+1}} h\right)^t \end{gathered}$$

and we have our result.

2.1.2 Consequences of PL

Most immediately, we will prove Brunn Minkowski on $\mathbb{R}^n$.

Brunn-Minkowski Proof

Given A, B Borel subsets of $\mathbb{R}^n$ it is immediate from definitions that

$$\mathbb{1}_{(1-t)A+tB} \geq (\mathbb{1}_A)^{1-t} (\mathbb{1}_B)^t$$

Applying PL,

$$\int_{\mathbb{R}^n} \mathbb{1}_{(1-t)A+tB} \geq \left(\int_{\mathbb{R}^n} \mathbb{1}_A\right)^{1-t} \left(\int_{\mathbb{R}^n} \mathbb{1}_B\right)^t \tag{2.13}$$

$$|(1-t)A + tB| \geq |A|^{1-t}|B|^t. \tag{2.14}$$

Which as we have shown, is the Brunn-Minkowski inequality.

Log-concave Measures

Before developing the theory of α-concave measures systematically, we will consider a measure μ to be log-concave if it is absolutely continuous with respect to the Lebesgue measure and its density function ρ is log-concave in the sense that $\log(\rho(x))$ is concave on the support of ρ in the usual manner.[1]

Theorem 2.1.3 (κ-Prékopa-Liendler) *Given μ a log κ-concave measure on $\mathbb{R}^n$, in the sense that $\frac{d\mu}{dx} = e^{-V}$ with $V((1-t)x+ty) \leq (1-\lambda)V(x) + \lambda V(y) - \kappa(1-\lambda)\lambda\frac{|x-y|^2}{2}$ holds for x, y in the support of μ and $t \in (0,1)$ and $\kappa \geq 0$. Then for f, g, h positive measurable functions on $\mathbb{R}^n$ and $t \in [0,1]$ such that*

$$f((1-t)x+ty) \geq e^{-\kappa(1-t)t\frac{|x-y|^2}{2}} g^{1-t}(x)h^t(y)$$

holds for all x, y in the support of μ. Then

$$\int f d\mu \geq \left(\int g d\mu\right)^{1-t} \left(\int h d\mu\right)^t.$$

[1] That is $\log(\rho((1-t)x+ty)) \geq (1-t)\log(\rho(x)) + t\log(\rho(y))$ for x, y such $\rho(x), \rho(y) > 0$. If one is willing to consider concavity on the extended real line, there is no need for the restriction of the domain to the support. Equivalently $\rho((1-t)x+ty) \geq \rho^{1-t}(x)\rho^t(y)$.

Proof For x, y both belonging to the support of μ,

$$
\begin{aligned}
f\frac{d\mu}{dx}(1-t)x+ty) &= f((1-t)x+ty)e^{-V((1-t)x+ty)} \\
&\geq e^{-\kappa(1-t)t\frac{|x-y|^2}{2}}\left(g^{1-t}(x)h^t(y)\right)e^{-(1-t)V(x)}e^{-tV(y)}e^{\kappa(1-t)t\frac{|x-y|^2}{2}} \\
&= (g\frac{d\mu}{dx})^{1-t}(x)(h\frac{d\mu}{dx})^t(y).
\end{aligned}
$$

If either x or y is outside of the support of μ, the inequality above holds trivially. Hence we may apply PL for the Lebesgue measure,

$$
\begin{aligned}
\int f d\mu &= \int f(z)\frac{d\mu}{dz}dz && (2.15) \\
&\geq \left(\int g(z)\frac{d\mu}{dz}dz\right)^{1-t}\left(\int h(z)\frac{d\mu}{dz}dz\right)^t && (2.16) \\
&= \left(\int g d\mu\right)^{1-t}\left(\int h d\mu\right)^t. && (2.17)
\end{aligned}
$$

Corollary 2.1.4 *If μ is a Log κ-concave measure, it satisfies a curved Brunn-Minkowski in the following sense*

$$
\mu((1-t)A+tB) \geq e^{\kappa(1-t)t\frac{d^2(A,B)}{2}}\mu^{1-t}(A)\mu^t(B).
$$

Proof Apply theorem 2.1.3 to the functions

$$
\begin{aligned}
f &= e^{-\kappa t(1-t)\frac{d^2(A,B)}{2}}\mathbb{1}_{(1-t)A+tB} \\
g &= \mathbb{1}_A \\
h &= \mathbb{1}_B.
\end{aligned}
$$

When $\kappa = 0$ the inequality reduces to the case of an ordinary log-concave measure, and shows that log-concavity of a measure is equivalent to satisfying a Prékopa-Leindler inequality.

2.1.3 Isoperimetry for $\mathbb{R}^n$

The classical isoperimetric inequality in Euclidean space, answers the question: given a fixed surface area, what is the maximal volume it can enclose? Or, equivalently, for a shape of fixed volume, what is the minimal surface area one can attain. Since antiquity, it has been more or less known that the answer should be that the Euclidean ball should provide minimizing conditions. However, a proof of this intuition was much more elusive. With the Brunn-Minkowski inequality in hand, we can construct a simple proof.

We first introduce the notion of a (general) surface measure. In a metric measure space (X, μ), for $A \subset X$ we define

$$A^h = \{x \in X : \exists a \in A, d(x, a) < h\}.$$

Notice A^h is open and hence Borel. We can define the surface measure of A as

$$\mu^+(A) = \liminf_{h \to 0} \frac{\mu(A^h) - \mu(A)}{h}.$$

For $A \subset \mathbb{R}^n$ with $c = |A| = |B_R|$, where B_R is the centered sphere of radius R chosen to attain equality. Now notice that $A^h = A + B_h$ so that by Brunn-Minkowski, $|A^h| \geq (|A|^{\frac{1}{n}} + |B_h|^{\frac{1}{n}})^n$. Applying this to the numerator of the surface measure quotient of A and rephrasing in terms of the sphere we have the following,

$$|A^h| - |A| \geq (|A|^{\frac{1}{n}} + |B_h|^{\frac{1}{n}})^n - |A| \tag{2.18}$$

$$= |B_1|((R + h)^n - R). \tag{2.19}$$

Thus it follows by the above, and computing directly the surface measure of R-ball.

$$|A|^+_{n-1} \geq n|B_1|R^{n-1} = |B_R|^+_{n-1}$$

Thus we have

$$|A|^+_{n-1} \geq \frac{n|A|^{\frac{n-1}{n}} \pi^{\frac{1}{2}}}{\Gamma^{\frac{1}{n}}(1 + \frac{n}{2})}.$$

2.2 The Localization Lemma

The localization lemma of Lovasz and Simonovits originally published in [32] as a means to an improved volume algorithm for convex bodies. We state the result now, though it may not be clear the results relation to Brunn-Minkowski.

Theorem 2.2.1 *If f, g are integrable lower semi continuous functions from $\mathbb{R}^n$ to $\mathbb{R}$ such that*

$$\int f dx > 0, \int g dx > 0, \tag{2.20}$$

Then there exists points $a, b \in \mathbb{R}^n$ a positive, affine function l on $(0,1)$ such that

$$\int_0^1 l^{n-1}(t) f(ta + (1-t)b) dx > 0, \int_0^1 l^{n-1}(t) g(ta + (1-t)b) dx > 0. \tag{2.21}$$

The technique in the proof is a bisection method that dates back at least to Payne and Weinburger [35]. The lemma was recognized as a powerful tool towards certain integral relations in $\mathbb{R}^n$. It allows reduction to dimension $n = 1$, although in a form of different one dimensional relations. It has found numerous applications in different problems of multidimensional Analysis and Geometry, such as isoperimetric problems over convex bodies, log-concave and more general hyperbolic measures, as well as Khinchine and dilation-type inequalities for examples see [5, 7, 8, 10–13, 21, 27, 29, 38].

Chapter 3

α-Concave Measures

3.1 Gaussian Measures

The standard Gaussian Measure γ_n on $\mathbb{R}^n$ can be given by the density function φ_n

$$\varphi_n(x) = (2\pi)^{-\frac{n}{2}} e^{-|x|^2/2}. \tag{3.1}$$

With the notation $\Phi(t) = \int_{-\infty}^{t} (2\pi)^{-\frac{2}{2}} e^{-x^2/2} dx$ we write the following

Theorem 3.1.1 *For A measurable in $\mathbb{R}^N$ and $h \in [0, \infty)$*

$$\gamma_n(A_h) \geq \Phi(\Phi^{-1}(\gamma_n(A)) + h). \tag{3.2}$$

This is the statement that for a fixed volume the space that minimizes surface area are the half-spaces.

Theorem 3.1.2 *[6, 10] Suppose that μ is a measure on $\mathbb{R}^n$ that has a log-concave density with respect to the standard Gaussian. Then*

$$\mu(A_h) \geq \Phi(\Phi^{-1}(\mu(A)) + h). \tag{3.3}$$

It is a result of Caffarelli [20] using optimal transport theory, that the measures of the above type are actually Lipschitz pushfowards of the Gaussian. However within our current frame work, the result above can be proved via elementary means. We will show that the Lovász and Simmonovits localization lemma reduces the above theorem to the one dimensional result. The details of the one dimensiona result are elementary and can be found in [6, 10].

Proof By approximation, one can see that if the result fails, there will exist A open $p > 0$ such that $\mu(A) > p$ while, $\mu\left(\overline{A_h}\right) < \Phi(\Phi^{-1}(p) + h)$. Writing the density of μ as $\rho\varphi_n$ where ρ is a log-concave function. Then defining

$$f = (\mathbb{1}_A - p)\rho\varphi_n \quad g = \left(\Phi(\Phi^{-1}(p) + h) - \mathbb{1}_{\overline{A_h}}\right)\rho\varphi_n$$

By construction

$$\int f dx = \mu(A) - p > 0 \qquad \int g dx = \Phi(\Phi^{-1}(p) + h) - \mu\left(\overline{A_h}\right) > 0.$$

By Lovász and Simmonovits there exists $a, b \in \mathbb{R}^n$ and $\ell(t)$, a log-concave function[1] on $[0, 1]$ such that

$$\int_0^1 f((a + t(b - a))\ell(t)dt > 0 \qquad \int_0^1 g((a + t(b - a))\ell(t)dt > 0$$

Notice that we can exclude the case that $a = b$, as this would require a belonging to the obviously empty $A \setminus A_h$. Now using the change of variables $z = t|b - a|$, writing $r = |b - a|$, $\theta = (b - a)/r$ and we have

$$\int_0^r (\mathbb{1}_A(a + z\theta) - p)\frac{\varphi_n(a + z\theta)}{\varphi_1(z)}\rho(a + z\theta)\ell(z/r)\varphi_1(z)dz > 0$$
$$\int_0^r (\Phi(\Phi^{-1}(p) + h) - \mathbb{1}_{\overline{A_h}}(a + z\theta))\frac{\varphi(a + z\theta)}{\varphi_1(z)}\rho(a + z\theta)\ell(z/r)\varphi_1(z)dz > 0.$$

The probability measure ν induced by the density $\frac{\varphi_n(a+z\theta}{\varphi_1(z)}\rho(a+z\theta)\ell(z/r)\varphi_1(z)$ is log-concave with respect to φ_1. This follows since $\frac{\varphi_n(a+z\theta)}{\varphi_1(z)}$ is log-concave (due to $|\theta| = 1$), and the product of log-concave functions is again log-concave. Taking

$$B = \{z : a + z\theta \in A\}$$
$$C = \{z : a + z\theta \in \overline{A_h}\},$$

we can rewrite the above conclusion equivalently as

$$\int \mathbb{1}_B - p d\nu > 0 \qquad \int \Phi(\Phi^{-1}(p) + h) - \mathbb{1}_C d\nu.$$

[1] The localization lemma actually guarantees us a function whose $\frac{1}{n-1}$th power is affine, a statement stronger than just log-concavity

That is, $\nu(B) > p$ and $\nu(C) < \Phi(\Phi^{-1}(p) + h)$. But since $B_h \subseteq C$, we have

$$\begin{aligned}\nu(B^h) &\leq \nu(C)\\ &< \Phi(\Phi^{-1}(p) + h)\\ &\leq \Phi(\Phi^{-1}(\nu(B)) + h).\end{aligned}$$

Thus it suffices to prove the one dimensional result see [6, 10].

Also as a direct consequence of κ Prékopa Leindler (Theorem 2.1.3) we have the following,

Theorem 3.1.3

$$\gamma_n((1-t)A + tB) \geq e^{(1-t)t\frac{d^2(A,B)}{2}} \gamma_n^{1-t}(A)\gamma_n^t(B)$$

Proof Computing directly, $Hess(V) = Hess(|x|^2/2) = Id$ shows that the Gaussian is log 1-concave. Applying theorem 2.1.3, we have our result.

3.2 α-Concave Measures

Definition For a complete, locally convex space E we will call a Radon measure μ, α-concave, if $t \in [0, 1]$, $A, B \subset E$ measurable implies

$$\mu_*((1-t)A + tB) \geq ((1-t)\mu^\alpha(A) + t\mu^\alpha(B))^{\frac{1}{\alpha}}. \tag{3.4}$$

Here μ_* represents the inner-measure taken as the supremum of compact subsets.

Theorem 3.2.1 *When μ is α concave, the supp(μ) is a convex set.*

Proof Given $x, y \in supp(\mu)$, and U open neighborhood of $(1-t)x + ty$, by continuity of vector space operations, there exists open neighborhoods of x and y denoted U_x and U_y such that $(1-t)U_x + tU_y \subset U$. Thus

$$\begin{aligned}\mu(U) &\geq \mu((1-t)U_x + tU_y)\\ &\geq \min(\mu(U_x), \mu(U_y)) > 0.\end{aligned}$$

3.2.1 Support, dimension and characterizations

The support $H_\mu = \mathrm{supp}(\mu)$ of any Radon measure μ on E is defined as the smallest closed subset of E of full measure, so that $\mu(E \setminus H_\mu) = 0$. If μ is hyperbolic, then the set H_μ is necessarily convex, as follows from (1.4). This set has some dimension

$$k = \dim(\mu) = \dim(H_\mu),$$

finite or not, which is called the dimension of the hyperbolic measure μ. If it is finite, absolute continuity of μ will always be understood with respect to the k-dimensional Lebesgue measure on H_μ.

First, let us recall an important general property of hyperbolic measures proven by Borell.

Theorem 3.2.2 *[16]. If μ is a hyperbolic probability measure on a locally convex space E, then for any additive subgroup H of E, ether $\mu_*(H) = 0$ or $\mu_*(H) = 1$.*

In particular, any μ-measurable affine subspace of E has measure either zero or one.

In [16, 17], Borell also gave a full description of α-concave measures. Similarly to (3.4), a non-negative function f defined on a convex subset H of E is called β-concave, if it satisfies

$$f\big((1-t)x + ty\big) \geq \Big[(1-t)f(x)^\beta + tf(y)^\beta\Big]^{1/\beta} \tag{3.5}$$

for all $t \in (0,1)$ and all points $x, y \in H$ such that $f(x) > 0$ and $f(y) > 0$. The right-hand side is understood in the usual limit sense for the values $\beta = -\infty$, $\beta = 0$ and $\beta = \infty$.

Theorem 3.2.3 *[16] . If μ is a finite α-concave measure on $\mathbb{R}^n$ of dimension $k = \dim(\mu)$, then $\alpha \leq \frac{1}{k}$. Moreover, μ is absolutely continuous with respect to Lebesgue measure on H_μ and has density f which is positive, finite, and β-concave on the relative interior of H_μ, where*

$$\beta = \frac{\alpha}{1 - \alpha k}.$$

Conversely, if a measure μ on $\mathbb{R}^n$ is supported on a convex set H of dimension k and has there a positive, β-concave density f with $\beta \geq -\frac{1}{k}$, then μ is α-concave.

Note that β is continuously increasing in the range $[-\frac{1}{k}, \infty]$, when α is varying in $[-\infty, \frac{1}{k}]$.

In the extremal case $\alpha = \frac{1}{k}$, the density $f(x) = \frac{d\mu(x)}{dx}$ is ∞-concave and is therefore constant: Up to a factor, μ must be the k-dimensional Lebesgue measure on H_μ.

More generally, if $\alpha \leq \frac{1}{k}$, $\alpha \neq 0$, the density has the form

$$f(x) = V(x)^{\frac{1}{\alpha} - k}$$

for some function $V : \Omega \to (0, \infty)$ on the relative interior Ω of H_μ, which is concave in case $\alpha > 0$, and is convex in case $\alpha < 0$. In particular, the formula

$$f(x) = V(x)^{-k}$$

describes all k-dimensional hyperbolic measures ($\alpha = -\infty$). If $\alpha = 0$, then necessarily $f(x) = e^{-V(x)}$ for some convex function $V : \Omega \to \mathbb{R}$.

As for general locally convex spaces, another theorem due to Borell reduces the question to Theorem 3.2.3.

Theorem 3.2.4 *[16]. A Radon probability measure μ on the locally convex space E is α-concave, if and only if the image of μ under any linear continuous map $T : E \to \mathbb{R}^n$ is an α-concave measure on $\mathbb{R}^n$.*

For special spaces in this characterization one may consider linear continuous maps T from a sufficiently rich family. For example, when $E = C[0, 1]$ is the Banach space of all continuous functions on $[0, 1]$ with the maximum-norm, the measure μ is α-concave, if and only if the image of μ under any map of the form

$$Tx = (x(t_1), \ldots, x(t_n)), \qquad x \in C[0, 1], \quad t_1, \ldots, t_n \in [0, 1],$$

is an α-concave measure on $\mathbb{R}^n$. Similarly, when $E = \mathbb{R}^\infty$ is the space of all sequences of real numbers (with the product topology), it is sufficient to consider the standard projections

$$T_n x = (x_1, \ldots, x_n), \qquad x = (x_1, \ldots, x_n, \ldots) \in \mathbb{R}^\infty. \tag{3.6}$$

The next general observation emphasizes that infinite dimensional α-concave measures may not have a positive parameter of convexity. Apparently, it was not stated explicitly in the literature, so we include the proof. As usual, E' denotes the dual spaces of all linear continuous functionals on E.

Theorem 3.2.5 *For $\alpha > 0$, any α-concave finite measure μ on a locally convex space E has finite dimension and is compactly supported.*

Proof First, suppose to the contrary that μ is infinite dimensional. We may assume that $H_\mu = \mathrm{supp}(\mu)$ contains the origin. Since H_μ is not contained in any finite dimensional subspace of E, for each n, one can find linearly independent vectors $v_1, \dots, v_n \in H_\mu$. Each point $x \in E$ has a representation $x = c_1(x)v_1 + \dots + c_n(x)v_n + y$ with some $c_i \in E'$, where $y = y(x)$ is linearly independent of all v_i (cf. [37], Lemma 4.21). Consider the linear map $T(x) = (c_1(x), \dots, c_n(x))$, which is continuously acting from E to $\mathbb{R}^n$. Then the image $\nu = \mu T^{-1}$ of μ is a finite α-concave measure on $\mathbb{R}^n$.

Let us see that ν is full dimensional. Otherwise, ν is supported on some hyperplane in $\mathbb{R}^n$ described by the equation $a_1 y_1 + \dots + a_n y_n = a_0$, where the coefficients $a_i \in \mathbb{R}$ are not all zero. Moreover, since $0 \in H_\mu$, any neighborhood of 0 has a positive μ-measure, so

$$\mu\{x \in E : |T(x)| < \varepsilon\} > 0,$$

for any $\varepsilon > 0$. Hence, necessarily $a_0 = 0$. This implies that μ is supported on the closed linear subspace H of E described by the equation $a_1 c_1(x) + \dots + a_n c_n(x) = 0$. Here, at least one of the coefficient, say a_i, is non-zero. Since $c_i(v_i) = 1 \neq 0$, we obtain that $v_i \notin H$. But this would mean that $H_\mu \cap H$ is a proper closed subset of the support of μ, while H_μ has a full μ-measure, a contradiction.

Hence, $\dim(\nu) = n$. By Theorem 3.2.3, this gives $\alpha \leq \frac{1}{n}$, and since n was arbitrary, we conclude that $\alpha \leq 0$ which contradicts to the hypothesis $\alpha > 0$.

Thus, μ must be supported on a finite dimensional affine subspace $H \subset E$. To prove compactness of the support, we may assume that $H = E = \mathbb{R}^n$ and $\dim(\mu) = n$. Then, μ is supported on an open convex set $\Omega \subset \mathbb{R}^n$, where it has density of the form

$$f(x) = V(x)^\gamma, \qquad \gamma = \frac{1}{\alpha} - n \quad \big(0 < \alpha \leq \frac{1}{n}\big),$$

for some concave function $V : \Omega \to (0, \infty)$. The case $\gamma = 0$ is possible, but then $f(x) = c$ for some constant $c > 0$, which implies $\mu(\mathbb{R}^n) = c\,|\Omega|$. Since μ is finite, Ω has to be bounded, and so $H_\mu = \text{clos}(\Omega)$ is compact.

Now, let $\gamma > 0$. Suppose that Ω is unbounded (to justify several notations below). It is known (cf. e.g. [11]) that $f(x) \to 0$ uniformly as $|x| \to \infty$ $(x \in \Omega)$. In particular, f is bounded, that is, $A = \sup_{x \in \Omega} f(x)$ is finite. Here, we may assume that the sup is asymptotically attained at $x_0 = 0$ for some sequence $x_l \to 0$, $x_l \in \Omega$. Choose $r > 0$ so that $f(x) < \frac{1}{2} A$ or $V(x) < (\frac{1}{2} A)^{1/\gamma}$ whenever $|x| \geq r$ and $x \in \Omega$. For such x, the sequence

$$\lambda_l(x) = \sup\{\lambda > 1 : x_l + \lambda(x - x_l) \in \Omega\}$$

has a limit $\lambda(x) = \sup\{\lambda : \lambda_l \in \Omega\} > 1$. Consider the convex functions

$$\psi_l(\lambda) = V(x_l) - V(x_l + \lambda(x - x_l)), \qquad 0 \leq \lambda < \lambda_l(x).$$

We have $\psi_l(0) = 0$ and $\psi_l(1) = V(x_l) - V(x) > C = A^{1/\gamma}(1 - 2^{-1/\gamma})$, for all k large enough. Hence, $\psi_l(\lambda) \geq C\lambda$, for all admissible $\lambda \geq 1$, and letting $l \to \infty$, we obtain

$$C\lambda \leq A^{1/\gamma} - V(\lambda x), \qquad 1 \leq \lambda < \lambda(x) \ \ (|x| \geq r,\ x \in \Omega).$$

But V is non-negative, so necessarily $\lambda(x) \leq \frac{1}{1 - 2^{-1/\gamma}}$. This proves boundedness of Ω.

3.2.2 Examples

1. The normalized Lebesgue measure on every convex body $K \subset \mathbb{R}^n$ is $\frac{1}{n}$-concave.

2. Any Gaussian measure on a locally convex space E is log-concave. In particular, the Wiener measure on $C[0, 1]$ is such.

3. The standard Cauchy measure μ_1 on $\mathbb{R}$ with density $f(x) = \frac{1}{\pi(1+x^2)}$ is α-concave with $\alpha = -1$ (which is optimal). More generally, the n-dimensional Cauchy measure μ_n on $\mathbb{R}^n$ with density

$$f_n(x) = \frac{c_n}{(1 + |x|^2)^{(n+1)/2}}$$

is (-1)-concave (c_n is a normalizing constant so that μ_n is probability).

4. Although the above density f_n essentially depends on the dimension, the measure μ_n has a dimension-free essense. All marginals of μ_n coincide with μ_1 and moreover, there is a unique Borel probability measure μ on $\mathbb{R}^\infty$ (an infinite dimensional

Cauchy measure) which is pushed forward to μ_n by the standard projection T_n from (2.2). This measure can also be introduced as the distribution of the random sequence

$$X = \Big(\frac{Z_1}{\zeta}, \frac{Z_2}{\zeta}, \dots\Big),$$

where the random variables $\zeta, Z_1, Z_2, \dots$ are independent and all have a standard normal distribution. Thus, μ is (-1)-concave on $\mathbb{R}^\infty$.

5. This example is mentioned in [16]. Given $d > 0$ (real), let χ_d be a positive random variable such that χ_d^2 has the χ^2-distribution with d degrees of freedom, i.e., with density

$$f_d(r) = \frac{1}{2^{d/2}\Gamma(d/2)}\, r^{d/2-1}\, e^{-r/2}, \qquad r > 0.$$

Let W be the standard Wiener process (independent of χ_d) viewed as a random function in $C[0,1]$. Then the random function

$$X(t) = \frac{\sqrt{d}}{\chi_d}\, W(t), \quad t \in [0,1],$$

has the distribution μ which is α-concave on $C[0,1]$ with $\alpha = -\frac{1}{d}$. It is called the Student measure (and also Cauchy in case $d = 1$ similarly to the previous example).

Chapter 4

Localization

Theorem 4.0.6 *Let μ be a finite α-concave measure on a complete locally convex space E, and let $u, v : E \to \mathbb{R}$ be lower semi-continuous μ-integrable functions such that*

$$\int_E u\, d\mu > 0, \qquad \int_E v\, d\mu > 0. \tag{4.1}$$

Then, for some points $a, b \in E$ and some finite α-concave measure ν supported on the segment $\Delta = [a, b]$,

$$\int_\Delta u\, d\nu > 0, \qquad \int_\Delta v\, d\nu > 0. \tag{4.2}$$

Note that lower semicontinuous functions are bounded below on any compact set. Hence, their integrals over compactly supported finite measures such as (4.1) and (4.2) always exist. As an example, the indicator functions of open subsets of E are all lower semicontinuous.

The completeness assumption (meaning that every Cauchy net in E is convergent) is quite natural. It ensures that the closed convex hull of any compact set in E is also compact. In that case any finite Radon measure μ on E has a stronger property

$$\sup\{\mu(K) : K \subset E \text{ convex compact}\} = \mu(E). \tag{4.3}$$

This property is crucial in some applications, but without completeness it is not true in general. (Its validity remains unclear e.g. for Radon Gaussian measures.)

One can also give a geometric variant of Theorem 2.2.1 together with a finer formulation of Theorem 4.0.6 in terms of extreme points of the set $\mathcal{P}_\alpha(u)$ of all α-concave probability measures supported on a convex compact set $K \subset E$ and such that $\int u\, d\mu \geq 0$

(for a continuous function u on K). As we already mentioned, this interesting approach to localization was developed by Fradelizi and Guédon [22]. It was shown there that in case $E = \mathbb{R}^n$ and $\alpha \leq \frac{1}{2}$, any extreme point in $\mathcal{P}_\alpha(u)$ is either a mass point or it is supported on an interval $\Delta \subset K$ with density $l^{(1-\alpha)/\alpha}$ (where l is a non-negative affine function on Δ). As will be explaned in Section 3, this property extends to general locally convex spaces, and then it easily implies Theorem 4.0.6.

One interesting application of Theorem 4.0.6 may be stated in terms of the following operation proposed in [38]. Given a Borel subset A in a closed convex set $F \subset E$ and a number $\delta \in [0,1]$, define

$$A_\delta = \Big\{ x \in A : m_\Delta(A) \geq 1 - \delta \ \text{ for any interval } \Delta \subset F \text{ such that } x \in \Delta \Big\},$$

where m_Δ denotes the normalized one-dimensional Lebesgue measure on Δ.

For example, if $F = E$ and A is the complement to a centrally symmetric, open, convex set $B \subset E$, then $A_\delta = E \setminus (\frac{2}{\delta} - 1)B$ represents the complement to the corresponding dilation of B.

Theorem 4.0.7 *Let μ be an α-concave probability measure on a complete locally convex space E supported on a closed convex set F $(-\infty < \alpha \leq 1)$. For any Borel set A in F and for all $\delta \in [0,1]$ such that $\mu^*(A_\delta) > 0$,*

$$\mu(A) \ \geq\ \big[\, \delta \mu^*(A_\delta)^\alpha + (1-\delta) \big]^{1/\alpha}. \tag{4.4}$$

Here μ^* denotes the outer measure (which is not needed, when E is a Fréchet space). This relation resembles very much the definition (3.4).

In the important particular case $\alpha = 0$ (i.e., for log-concave measures), (4.4) becomes

$$\mu(A) \ \geq\ \mu^*(A_\delta)^\delta.$$

It was discovered by Nazarov, Sodin and Vol'berg [38]. The extension of this result to the class of α-concave measures in the form (4.4) is settled in [5] and [21], still for finite dimensional spaces. All proofs are essentially based on Theorem 2.2.1 or its modifications to reduce (4.4) to dimension one (although the one dimensional case appears to be rather delicate). Here we make another step removing the dimensionality of the space assumption, cf. Section 6.

We will develop extensions of Fradelizi-Guédon's theorem and Lovász-Simonovits' bisection argument. In particular, the existence of needles which we understand in a somewhat weaker sense is proved for probability measures on Fréchet spaces that satisfy the zero-one law. This can be used as an approach towards Theorems 2.2.1 -4.0.6, but potenitally may have a wider range of applications.

We do not try to describe in detail results and techniques in dimension one, but mainly focus on their extensions to the setting of infinite dimensional spaces.

4.1 Extreme α-concave measures

Given a convex compact set K in a locally convex space E, denote by $\mathcal{M}_\alpha(K)$ the collection of all α-concave probability measures with support contained in K. For a continuous function u on K, we consider the subcollection

$$\mathcal{P}_\alpha(u) = \left\{ \mu \in \mathcal{M}_\alpha(K) : \int u \, d\mu \geq 0 \right\}$$

together with its closed convex hull $\widetilde{\mathcal{P}}_\alpha(u)$ in the locally convex space $\mathcal{M}(K)$ of all signed Radon measures on K endowed with the topology of weak convergence. The latter space is dual to the space $C(K)$ of all continuous functions on K, and $\widetilde{\mathcal{P}}_\alpha(u)$ is a convex compact subset of $\mathcal{M}(K)$.

What are extreme points of $\widetilde{\mathcal{P}}_\alpha(u)$? Using a general theorem due to D. P. Milman, one can only say that all such points lie in $\mathcal{P}_\alpha(u)$ (cf. [14], p.124, or [36] for a detail discussion of Krein-Milman's theorem). A full answer to this question is given in Fradelizi-Guédon's theorem, which we formulate below in the setting of abstract locally convex spaces.

Theorem 4.1.1 *Given a continuous function u on K and $-\infty \leq \alpha \leq 1$, any extreme point μ in $\widetilde{\mathcal{P}}_\alpha(u)$ has the dimension $\dim(\mu) \leq 1$. Moreover, in case $\alpha \leq \frac{1}{2}$,*

1) *μ is either a mass point at $x \in K$ such that $u(x) \geq 0$; or*
2) *μ is supported on an interval $\Delta = [a, b] \subset K$ with density*

$$\frac{d\mu(x)}{dm_\Delta(x)} = l(x)^{(1-\alpha)/\alpha} \tag{4.1}$$

with respect to the uniform measure m_Δ, where l is a non-negative affine function on Δ such that $\int_a^x u\, d\mu > 0$ and $\int_x^b u\, d\mu > 0$, for all $x \in (a,b)$.

In particular, any α-concave probability measure supported on K, belongs to the closed convex hull of the family of all one-dimensional α-concave probability measures supported on K having density of the form (4.1).

We only consider the first assertion of the theorem. The second part is a purely one dimensional statement, and we refer to [22].

Proof Suppose that a measure $\mu \in \mathcal{P}_\alpha(u)$ has the dimension $\dim(\mu) \geq 2$. For simplicity, let the origin belong to the relative interior G of the support H_μ of μ. Then one may find linearly independent vectors x and y such that $\pm x$ and $\pm y$ are all in G. On the linear hull $L(x,y)$ of x and y (which is a 2-dimensional linear subspace of E), define linear functionals λ_x and λ_y by putting

$$\lambda_x(x) = \lambda_y(y) = 1,$$

$$\lambda_x(y) = \lambda_y(x) = 0.$$

They are continuous, so by the Hahn-Banach theorem, these functionals may be extended from $L(x,y)$ to the whole space E keeping linearity and continuity.

With these extended functionals, we can associate $\Lambda_\theta = \theta_1\lambda_x + \theta_2\lambda_y$, where $\theta = (\theta_1, \theta_2) \in \mathbb{S}^1$ (vectors on the unit sphere of $\mathbb{R}^2$). Note that these functionals are uniformly bounded on K, i.e.,

$$\sup_\theta \sup_{z\in K} |\Lambda_\theta(z)| \leq \sup_{z\in K} |\lambda_x(z)| + \sup_{z\in K} |\lambda_y(z)| < \infty. \tag{4.2}$$

Now, following in essence an argument of [22], define the map $\Phi : \mathbb{S}^1 \to \mathbb{R}$ by

$$\Phi(\theta) = \int_{\{\Lambda_\theta \geq 0\}} u\, d\mu.$$

By the construction, the set $\{\Lambda_\theta = 0\} \cap H_\mu$ represents a proper closed affine subspace of H_μ. So, $\mu\{\Lambda_\theta = 0\} = 0$ according to Theorem 4.1.1 (the zero-one law for hyperbolic measures). Hence, using (4.2), we may conclude that the map Φ is continuous.

In addition, we have the identity $\Phi(\theta) + \Phi(-\theta) = \int u\,d\mu$. Hence, the intermediate value theorem implies that there exists θ such that with $H_\theta^+ = \{\Lambda_\theta \geq 0\}$ and $H_\theta^- = \{\Lambda_\theta \leq 0\}$, we have

$$\int_{H_\theta^+} u\,d\mu = \int_{H_\theta^-} u\,d\mu = \frac{1}{2}\int_E u\,d\mu.$$

Necessarily, $t = \mu(H_\theta^-) > 0$ and $\mu(H_\theta^+) > 0$. Defining α-concave probability measures

$$\mu_0(A) = \frac{\mu(A \cap H_\theta^+)}{\mu(H_\theta^+)}, \qquad \mu_1(A) = \frac{\mu(A \cap H_\theta^-)}{\mu(H_\theta^-)},$$

we arrive at the representation $\mu = (1-t)\mu_0 + t\mu_1$ which means that μ is not extreme. ∎

One can now return to Theorem 4.0.6.

Proof Due to the property (4.3), and by the assumption (4.1),

$$\int_K \min(u,c)\,d\mu > 0, \qquad \int_K \min(v,c)\,d\mu > 0,$$

for some convex compact set $K \subset E$ and a constant $c > 0$. Moreover, since the function $\min(u,c)$ is lower semicontinuous and bounded, while μ is Radon,

$$\int_K \min(u,c)\,d\mu \;=\; \sup_g \int g\,d\mu,$$

where the sup is taken over all continuous functions on K such that $g \leq \min(u,c)$ (cf. e.g. [33], Chapter 2, or [15], Chapter 7). A similar identity also holds for $\min(v,c)$. This allows us to reduce the statement of the theorem to the case where both u and v are continuous on K.

In the latter case, let $u_0 = u - \int_K u\,d\mu$. Consider the functional $T(\mu) = \int_K v\,d\mu$. It is linear and continuous on $\mathcal{M}(K)$, and therefore being restricted to $\mathcal{P}_\alpha(u_0)$ it attains maximum at one of the extreme points ν. Since $\mu \in \mathcal{P}_\alpha(u_0)$, we conclude that

$$\int_K u_0\,d\nu \geq 0, \quad T(\nu) \geq T(\mu),$$

so, $\int_K u\,d\nu > 0$ and $\int_K v\,d\nu > 0$ which is (1.6). It remains to apply Theorem 4.1.1. ∎

A similar argument, based also on the second part of Theorem 4.1.1, yields Theorem 2.2.1. Indeed, the n-dimensional integrals (2.20) can be restricted to a sufficently large closed ball $K \subset \mathbb{R}^n$. The normalized Lebesgue measure on K is α-concave with $\alpha = \frac{1}{n}$. Hence, the extreme points in $\mathcal{P}_\alpha(u)$ are at most one dimensional and have densities of the form l^{n-1} (if they are not Dirac measures).

4.2 Bisection and needles on Fréchet spaces

The notion of a needle was proposed by Lovász and Simonovits for the proof of Theorem 2.2.1 (Localization Lemma, cf. also [29]). Previously, it appeared implicitly in [35] and may be viewed as development of the Hadwiger-Ohmann bisection approach to the Brunn-Minkowski inequality ([19, 28], cf. also [26] for closely related ideas).

As shown in [31], starting from (2.20), one can construct a decreasing sequence of compact convex bodies K_l in $\mathbb{R}^n$ that are shrinking to some segment $\Delta = [a, b]$ and are such that, for each l,

$$\int_{K_l} u(x)\,dx > 0, \qquad \int_{K_l} v(x)\,dx > 0.$$

Moreover, choosing a further subsequence (if necessary) and applying the Brunn-Minkowski inequality in $\mathbb{R}^n$, one gets in the limit

$$\begin{aligned}
\lim_{l\to\infty} \frac{1}{|K_l|} \int_{K_l} u(x)\,dx &= \int_\Delta \psi^{n-1}(x)\,u(x)\,dx,\\
\lim_{l\to\infty} \frac{1}{|K_l|} \int_{K_l} v(x)\,dx &= \int_\Delta \psi^{n-1}(x)\,v(x)\,dx,
\end{aligned}$$

for some non-negative concave function ψ on Δ. Here $|K_l|$ denotes the n-dimensional volume, while the integration on the right-hand side is with respect to the linear Lebesgue measure on the segment. In this way, one may obtain a slightly weaker variant of (2.21) with ψ in place of l, and with non-strict inequalities. An additional argument of a similar flavour was then developed in [31] to make ψ affine (while the strict inequalities in (2.21) are easily achieved by applying the conclusion to functions $u - \varepsilon w$ and $v - \varepsilon w$, where $w > 0$ is integrable, continuous, and $\varepsilon > 0$ is small enough). The last step shows that for K_l one may take infinitesimal truncated cylinders with main axis Δ; it is in this sense the limit one dimensional measure $l^{n-1}(x)\,dx$ on Δ may be considered a needle.

The aim of this section is to extend this construction to the setting of Fréchet, i.e., complete metrizable locally convex spaces. For example, E may be a Banach space, but there also other important spaces that are not Banach, such as the space $E = \mathbb{R}^\infty$. Note that any finite Borel measure on a Fréchet space is Radon.

While one cannot speak about the Lebesgue measure when E is infinite dimensional, the main hypothesis (2.20) may readily be stated like (4.1) with integration with respect to a given (finite) Borel measure μ on E.

The space of all finite Borel measures on E is endowed with the topology of weak convergence. In particular, $\mu_l \to \mu$ (weakly), if and only if

$$\int u\, d\mu_l \to \int u\, d\mu \qquad (\text{as} \ \ l \to \infty)$$

for any bounded continuous functions u on E. As was noticed in [16], the class of all α-concave probability measures on E is closed in the weak topology.

Definition Let μ be a finite Borel measure on E. A Borel probability measure ν will be called a needle of μ, if it is supported on a segment $[a, b] \subset E$ and can be obtained as the weak limit of probability measures

$$\mu_l(A) = \frac{1}{\mu(K_l)}\, \mu(A \cap K_l), \qquad (A \text{ is Borel}),$$

where K_l is some decreasing sequence of convex compact sets in E of positive μ-measure such that $\cap_l K_l = [a, b]$.

Here, all μ_l represent normalized restrictions of μ to K_l. In particular, all needles of a given α-concave measure are α-concave, as well. We do not require that K_l be asymptotically close to infinitesimal truncated cylinders.

Definition One says that a Borel probability measure μ on E satisfies the zero-one law, if any μ-measurable affine subspace of E has μ-measure either 0 or 1.

For example, this important property holds true for all (Radon) Gaussian measures. More generally, it is satisfied by any hyperbolic probability measure, as follows from Borell's Theorem 3.2.2.

With these definitions, Theorem 4.0.6 admits the following refinement.

Theorem 4.2.1 *Suppose that a Borel probability measure μ on a Fréchet space E satisfies the zero-one law. Let $u, v : E \to \mathbb{R}$ be lower semi-continuous μ-integrable functions such that*

$$\int u \, d\mu > 0, \qquad \int v \, d\mu > 0.$$

Then, these inequalities also hold for some needle ν of μ. Moreover, if μ is supported on a closed convex set F, then ν may be chosen to be supported on F, as well.

First assume that E is a separable Banach space with norm $\|\cdot\|$, and let E' denote the dual space (of all linear continous functionals on E) with norm $\|\cdot\|_*$. Suppose that any proper closed affine subspace of E has μ-measure zero. In this case, for the proof of Theorem 4.2.1 we use the construction similar to the one from the proof of Theorem 4.1.1.

Given 3 affinely independent points x, y, z in E, define linear functionals λ_x and λ_y on the linear hull $L_z(x, y)$ of $x - z$ and $y - z$ (which is a 2-dimensional linear subspace of E), by putting

$$\lambda_x(x - z) = \lambda_y(y - z) = 1, \tag{4.3}$$

$$\lambda_x(y - z) = \lambda_y(x - z) = 0. \tag{4.4}$$

By the Hahn-Banach theorem, these functionals may be extended by linearity to the whole space E without increasing their norms. This will always be assumed below.

Lemma 4.2.2 *Let $\{(x_n, y_n, z_n)\}_{n\geq 1}$ be affinely independent points in the Banach space E such that $x_n \to x$, $y_n \to y$, $z_n \to z$, where x, y, z are also affinely independent. Then the corresponding linear functionals λ_{x_n} and λ_{y_n} have uniformly bounded norms, i.e.,*

$$\sup_{n\geq 1} \|\lambda_{x_n}\|_* < \infty, \qquad \sup_{n\geq 1} \|\lambda_{y_n}\|_* < \infty.$$

Proof Define the lines

$$\begin{aligned} L_z(x) &= \{z + r(x - z) : r \in \mathbb{R}\}, \\ L_z(y) &= \{z + r(y - z) : r \in \mathbb{R}\}. \end{aligned}$$

Then, for $w \in L_z(x,y)$, $\|w\| \leq 1$,

$$|\lambda_x(w)| \leq \mathrm{dist}^{-1}(x, L_z(y)), \quad |\lambda_y(w)| \leq \mathrm{dist}^{-1}(y, L_z(x)),$$

where we use the notation $\mathrm{dist}(w, A) = \inf\{\|w-a\| : a \in A\}$ (the shortest distance from a point to the set). The extended linear functionals should thus satisfy the above inequalities on the whole space E for all $\|w\| \leq 1$, i.e.,

$$\|\lambda_x\|_* \leq \mathrm{dist}^{-1}(x, L_z(y)), \quad \|\lambda_y\|_* \leq \mathrm{dist}^{-1}(y, L_z(x)). \tag{4.5}$$

Next, by shifting, one may assume that $z = 0$, in which case x and y are linearly independent and in particular $\|x\| > 0$ and $\|y\| > 0$. Using (4.5), it is enough to show that

$$\mathrm{dist}(x_n, L_{z_n}(y_n)) \geq c, \quad \text{for all } n \geq n_0,$$

with some n_0 and $c > 0$. Indeed, take an arbitrary point $w = z_n + r(y_n - z_n)$ in $L_{z_n}(y_n)$, $r \in \mathbb{R}$. By the triangle inequality,

$$\|x_n - w\| \geq |r|\,\|y_n - z_n\| - \|x_n - z_n\| \geq 2\|x_n - z_n\|,$$

where the last inequality holds whenever $|r| \geq 3\,\frac{\|x_n - z_n\|}{\|y_n - z_n\|}$. Hence, by the convergence assumption,

$$\|x_n - w\| \geq \|x\|, \qquad \text{for } \ |r| \geq r_0 = 4\,\frac{\|x\|}{\|y\|}, \ n \geq n_0.$$

In case $|r| \leq r_0$, again by the triangle inequality,

$$\begin{aligned} \|x_n - w\| \geq \quad & \|x - (z + ry)\| - \|x_n - x\| - |r|\,\|y_n - y\| - r\|z_n\| & (4.6) \\ \geq \quad & \mathrm{dist}(x, L_z(y)) - \|x_n - x\| - r_0\,\|y_n - y\| - r_0\|z_n\|. & (4.7) \end{aligned}$$

Here, the right-hand side is also separated from zero for sufficiently large n.

By a similar argument, $\mathrm{dist}(y_n, L_{z_n}(x_n)) \geq c$, for all $n \geq n_0$. ∎

Theorem 4.2.1 We begin with a series of reductions assuming without loss of generality that $\mu(E) = 1$.

Any Fréchet space with Radon probability measure μ has a subspace E_0 such that $\mu(E_0) = 1$, and in addition there exists a norm $\|\cdot\|$ on E_0 with respect to which E_0 is a separable reflexive Banach space whose closed balls are compact in E (see [15], Theorem 7.12.4).

In particular, all Borel subsets of E_0 are Borel in E. By the zero-one law (turning to a smaller subspace if necessary), we may assume that any proper affine subspace of E_0 which is closed for the topology of E_0 has measure zero. That is, for all $l \in E_0'$,

$$\mu\{l = c\} = 0, \quad c \in \mathbb{R}. \tag{4.8}$$

Second, it suffices to assume that the support of μ is compact, metrizable and convex. Indeed, by Ulam's theorem, there is an increasing sequence of compact sets $K_n \subset E_0$ such that $\mu(\cup_n K_n) = 1$. The closed convex hull of any compact set in E_0 is compact (which is true in any Banach and more generally complete locally convex spaces, cf. e.g. [30]). Therefore, all K_n may additionally be assumed to be convex. These sets will also be compact in E, so that the associated weak topologies in the spaces of Borel probability measures on K_n coincide, as well. By the dominated convergence theorem,

$$\lim_{n\to\infty} \int_{K_n} u\, d\mu = \int_E u\, d\mu, \qquad \lim_{n\to\infty} \int_{K_n} v\, d\mu = \int_E v\, d\mu,$$

so that $\int_{K_n} u\, d\mu > 0$ and $\int_{K_n} v\, d\mu > 0$ for large n. Hence, an application of the theorem to μ restricted and normalized to K_n would provide the desired one dimensional measure ν, a needle of μ_n and therefore of μ itself.

Thus, from now on, we may assume that E is a separable Banach space, and μ is a Borel probability measure on E which is supported on a convex compact set $K \subset E$ and is such that (4.8) holds true for all $l \in E'$.

We need only to prove the existence of ν such that $\int u\, d\nu \geq 0$ and $\int v\, d\nu \geq 0$. Since in this case we may apply the superficially weaker result to $u - \varepsilon$ and $v - \varepsilon$ for an $\varepsilon > 0$ chosen small enough to preserve the hypothesis.

In addition, it suffices to prove the result when u and v are both continuous. To see this, take u_n and v_n to be sequences continous functions increasing to lower semicontinuous u and v respectively. By the monotone convergence, $\lim_{n\to\infty} \int u_n\, d\mu = \int u\, d\mu > 0$ and $\lim_{n\to\infty} \int v_n\, d\mu = \int v\, d\mu > 0$, so we can take the approximating functions u_n and v_n to be such that $\int u_n\, d\mu > 0$ and $\int v_n\, d\mu > 0$. The theorem produces needles ν_n of μ supported on F and such that

$$\int u_n\, d\nu_n > 0, \qquad \int v_n\, d\nu_n > 0.$$

Since $u \geq u_n$ and $v \geq v_n$, every such measure ν_n will be the required needle.

Let us now turn to the construction procedure.

Given 3 affinely independent points x, y, z in E, consider the linear continuous functionals λ_x and λ_y on E introduced before Lemma 4.2.2 via the relations (4.3)-(4.4) and the Hahn-Banach theorem. To each point $\theta \in \mathbb{S}^1 = \{(t,s) : t^2+s^2 = 1\}$ we can associate a linear functional $\Lambda_\theta = t\lambda_x + s\lambda_y$ and define the function

$$\Psi : \mathbb{S}^1 \to \mathbb{R}, \qquad \theta = (t,s) \mapsto \int_{\{\ell_\theta(\xi - z) \geq 0\}} u(\xi)\, d\mu(\xi).$$

Since $\mu\{\xi : \Lambda_\theta(\xi - z) = 0\} = 0$ (cf. (4.4)), this function is continuous on $\mathbb{S}^1$. In addition, we have the identity

$$\Psi(-\theta) + \Psi(\theta) = \int_E u\, d\mu.$$

Hence, by the intermediate value theorem, there exists $\theta \in \mathbb{S}^1$ such that

$$\int_{\{\Lambda_\theta(\xi - z) \geq 0\}} u(\xi)\, d\mu(\xi) \;=\; \int_{\{\Lambda_\theta(\xi - z) \leq 0\}} u(\xi)\, d\mu(\xi) \;=\; \frac{1}{2} \int_E u\, d\mu.$$

Also,

$$\int_E v\, d\mu \;=\; \int_{\{\Lambda_\theta(\xi - z) \geq 0\}} v(\xi)\, d\mu(\xi) + \int_{\{\Lambda_\theta(\xi - z) \leq 0\}} v(\xi)\, d\mu(\xi) \;>\; 0,$$

so that at least one the last two integrals is positive. Let H^+ denote one of the hyperspaces $\{\Lambda_\theta(\xi - z) \geq 0\}$ or $\{\Lambda_\theta(\xi - z) \leq 0\}$ such that $\int_{H^+} v\, d\mu > 0$. Necessarily, $\mu(H^+) > 0$, and we may consider the normalized restriction μ^+ of μ to H^+ and will have the property that

$$\int_{H^+} u\, d\mu^+ > 0, \quad \int_{H^+} v\, d\mu^+ > 0. \tag{4.9}$$

This procedure can be performed step by step along a sequence $\{(x_n, y_n, z_n)\}_{n \geq 1}$ of affinely independent points, chosen to be dense in $K \times K \times K$. Let $\nu_1 = \mu^+$ be constructed according to the above procedure for (x_1, y_1, z_1) and with an associated point $\theta_1 = (t_1, s_1) \in \mathbb{S}^1$. Similarly, on the n-th step, given ν_n, let $\nu_{n+1} = \nu_n^+$ be constructed for the triple (x_n, y_n, z_n) and with the associated linear functional

$$\Lambda_{\theta_n} = \Lambda_{(t_n, s_n)} = t_n \lambda_{x_n} + s_n \lambda_{y_n}.$$

Since the space of all Borel probability measures on K is compact and metrizable for the weak topology, the sequence ν_n has a sub-sequential weak limit ν. In particular, from (4.9) we derive the desired property

$$\int_E u\, d\nu \geq 0, \qquad \int_E v\, d\nu \geq 0.$$

It remains to show that $\dim(H_\nu) \leq 1$. Suppose not, in this case there exists affinely independent x, y, z in the relative interior of H_ν that also contains the points $2z - x$ and $2z - y$. Without loss of generality, let $z = 0$, so that $\pm x$ and $\pm y$ belong to the relative interior of H_ν. By the density property, there exists a subsequence, say (x_k, y_k, z_k) such that $(x_k, y_k, z_k) \to (x, y, z)$.

By the construction, the measure ν_k^+ is supported on the half-space H_k^+, which is either $\{\xi : \Lambda_{(t_k,s_k)}(\xi - z_k) \geq 0\}$ or $\{\xi : \Lambda_{(t_k,s_k)}(\xi - z_k) \leq 0\}$. For definiteness, let it be the first half-space. Since all H_k^+ contain x and $-x$, we then have

$$\Lambda_{(t_k,s_k)}(x - z_k) \geq 0, \qquad \Lambda_{(t_k,s_k)}(-x - z_k) \geq 0, \tag{4.10}$$

and similarly for the point y.

Recall that by Lemma 4.2.2, we can obtain a uniform bound M such that

$$\|\Lambda_{(t_k,s_k)}\|_* \leq \|\lambda_{x_k}\|_* + \|\lambda_{y_k}\|_* \leq M \qquad \text{for all } k.$$

Hence, $\Lambda_{(t_k,s_k)}(z_k) \to 0$ and $\Lambda_{(t_k,s_k)}(x_k - x) \to 0$ as $k \to \infty$. But then by (4.6), necessarily $\Lambda_{(t_k,s_k)}(x_k) \to 0$, as well. By the same argument, $\Lambda_{(t_k,s_k)}(y_k) \to 0$.

On the other hand, according to the definition of $\Lambda_{(t_k,s_k)}$ via (4.3)-(4.4), for each k,

$$\Lambda_{(t_k,s_k)}(x_k - z_k) = t_k, \qquad \Lambda_{(t_k,s_k)}(y_k - z_k) = s_k,$$

thus implying that $\lim_{k\to\infty} t_k = \lim_{k\to\infty} s_k = 0$. But this is impossible since $t_k^2 + s_k^2 = 1$. This proves that $\dim(H_\nu) \leq 1$. ∎

4.3 Dilation and its properties

Before turning to Theorem 4.0.7, we first comment on the basic properties of the operation $A \to A_\delta$, where $\delta \in [0, 1]$ is viewed as parameter.

Let F be a closed convex subset of a locally convex space E with respect to which this operation is defined:

$$A_\delta = \Big\{x \in A : m_\Delta(A) \geq 1 - \delta, \text{ for any interval } \Delta \subset F \text{ such that } x \in \Delta\Big\}.$$

As before, m_Δ denotes a uniform distribution on Δ (understood as the Dirac measure, when the endpoints coincide). In this definiton, by the intervals Δ we mean closed intervals $[a, b]$ connecting arbitrary points a, b in F. Moreover, the requirement $x \in \Delta$ may equivalently be replaced by the condition that x is one of the endpoints of Δ.

Note that $A_1 = A$. If $0 \leq \delta < 1$, as an equivalent definition one could put

$$A_\delta = \Big\{x \in F : m_\Delta(A) \geq 1 - \delta, \text{ for any interval } \Delta \subset F \text{ such that } x \in \Delta\Big\}.$$

Indeed, in this case, if $x \in F \setminus A$, then $m_{[x,x]}(A) = 0 < 1 - \delta$ meaning that $x \notin A_\delta$ according to the second definition. Thus, for $\delta \in [0, 1)$, both definitions lead to the same set and we have the property $A_\delta \subset A$.

Lemma 4.3.1 *a) If $A \subset F$ is closed, then every set A_δ is closed as well.*

b) If E is a Fréchet space and A is Borel measurable in F, then every set A_δ is universally measurable.

Let us recall that a set in a Hausdorff topological space E is called universally measurable, if it belongs to the Lebesgue completion of the Borel σ-algebra with respect to an arbitrary Borel probability measure on E. In that case one may freely speak about the measures of these sets.

Proof For a Borel set A in F, consider the function

$$\psi(x, y) = \int 1_A \, dm_{[x,y]} = \int_0^1 1_A((1 - t)x + ty) \, dt, \qquad x, y \in F.$$

First assume that A is closed. Then, given a net $x_i \to x$, $y_i \to y$ in F indexed by a semi-ordered set I, we have

$$\limsup_{i \in I} 1_A((1 - t)x_i + ty_i) \leq 1_A((1 - t)x + ty).$$

After integration this implies

$$\limsup_{i\in I}\ \psi(x_i,y_i)\ \le\ \psi(x,y).$$

Indeed, the space $L^1[0,1]$ is separable, so the above relation is only to be checked for increasing sequences $i=i_n$ in I. But in that case one may apply the Lebesgue dominated convergence theorem. This means that ψ is upper semicontinuous on $F\times F$, and thus A_δ represents the intersection over all $y\in F$ of the closed sets $\{x\in A:\psi(x,y)\ge 1-\delta\}$.

In part b), assume that E is a Fréchet space. If A is Borel, then the function ψ is Borel measurable on $F\times F$, so the complement of A_δ in A,

$$A\setminus A_\delta=\{x\in A:\psi(x,y)<1-\delta,\ \text{for some } y\in A\},$$

represents the x-projection of a Borel set in $E\times E$. But every Borel set in a Polish space is Souslin, and therefore both $A\setminus A_\delta$ and A_δ are universally measurable (cf. [15], Corollary 6.6.7 and Theorem 7.4.1). ∎

There is an opposite operation representing a certain dilation or enlargement of sets. Given a Borel measurable set $B\subset F$ and $\delta\in[0,1)$, define

$$B^\delta\ =\ \bigcup_{m_\Delta(B)>\delta}\Delta\ =\ \Big\{x\in F: m_{[x,y]}(B)>\delta\ \text{ for some } y\in F\Big\}. \tag{4.11}$$

Here the union is running over all intervals $\Delta\subset F$ such that $m_\Delta(B)>\delta$.

Note that B^δ contains B (since all singletons in B participate in the above union).

Lemma 4.3.2 *For any $\delta\in[0,1)$ and any Borel set $B\subset F$, the complement $A=F\setminus B$ satisfies the dual relations*

$$F\setminus A_\delta=(F\setminus A)^\delta \qquad \textit{and} \qquad F\setminus B^\delta=(F\setminus B)_\delta.$$

In particular, B^δ is open in F, once B is open in F.

Proof Given $x\in F$, the property $x\notin A_\delta$ means that, for some interval $\Delta\subset F$ containing x, we have $m_\Delta(A)<1-\delta$, that is, $m_\Delta(B)>\delta$ meaning that $\Delta\subset B^\delta$. Therefore, $x\notin A_\delta\Leftrightarrow x\in B^\delta$. For the last assertion, it remains to recall Lemma 4.3.1. ∎

Lemma 4.3.3 *Let F be a convex closed set in E, and let T be a linear continuous map from E to another locally convex space E_1. For any Borel set $C \subset T(F)$,*

$$\big(T^{-1}(C) \cap F\big)^{\delta} = T^{-1}\big(C^{\delta}\big) \cap F,$$

where the operation $C \to C^{\delta}$ is understood with respect to the image $T(F)$.

Proof For all $a, b \in F$, the map T pushes forward the unform measure $m_{[a,b]}$ to $m_{[Ta,Tb]}$. Therefore, the pre-image $B = T^{-1}(C)$ has measure $m_{[a,b]}(B) = m_{[Ta,Tb]}(C)$, so

$$\big(B \cap F\big)^{\delta} \;=\; \bigcup_{m_{[Ta,Tb]}(C)>\delta} [a,b] \;=\; \bigcup_{m_{[x,y]}(C)>\delta} T^{-1}([x,y]) \cap F \;=\; T^{-1}\big(C^{\delta}\big) \cap F.$$

∎

When $E_1 = \mathbb{R}^n$ and C is a polytope, the dilated set C^{δ} is a polytope, as well. Hence, by Lemma 4.3.3, $(T^{-1}(C))^{\delta}$ represents the intersection of finitely many half-spaces.

4.4 The dual form and proof of Theorem 4.0.7

Following [5], let us reformulate Theorem 4.0.7 in terms of dilated sets. Putting $B = F \setminus A$ and using Lemma 4.3.2, the inequalty

$$\mu(A) \;\geq\; \big[\,\delta\mu^*(A_\delta)^{\alpha} + (1-\delta)\big]^{1/\alpha} \qquad (0 < \delta < 1) \tag{4.1}$$

is solved as $\mu_*(B^{\delta}) \geq R_\delta^{(\alpha)}(\mu(B))$, where

$$R_\delta^{(\alpha)}(p) = 1 - \left[\frac{(1-p)^{\alpha} - (1-\delta)}{\delta}\right]^{1/\alpha}. \tag{4.2}$$

More precisely, in the case $\alpha < 0$, the above expression is well-defined and represents a strictly concave, increasing function in $p \in [0,1]$. For $\alpha = 0$, it is understood in the limit sense as

$$R_\delta^{(0)}(p) = 1 - (1-p)^{1/\delta}, \qquad 0 \leq p \leq 1,$$

which is also strictly concave and increasing. In case $0 < \alpha \leq 1$, $R^{(\alpha)}(p)$ is defined to be (4.2) on the interval $0 \leq p \leq 1 - (1-\delta)^{1/\alpha}$ (when the expression makes sense) and we should put $R^{(\alpha)}(p) = 1$ on the remaining subinterval of $[0,1]$.

In all cases, $R_\delta^{(\alpha)} : [0,1] \to [0,1]$ represents a concave, continuous, non-decreasing function such that $R_\delta^{(\alpha)}(0) = 0$ and $R_\delta^{(\alpha)}(1) = 1$. Put $R_0^{(\alpha)}(p) = \lim_{\delta \downarrow 0} R_\delta^{(\alpha)}(p) = 1$ for $0 < p \le 1$ and $R_0^{(\alpha)}(0) = 0$.

Theorem 4.4.1 *Let μ be an α-concave probability measure on a complete locally convex space E supported on a convex closed set F $(-\infty < \alpha \le 1)$. For any Borel subset B of F and for all $\delta \in [0,1)$,*

$$\mu_*(B^\delta) \;\ge\; R_\delta^{(\alpha)}(\mu(B)). \tag{4.3}$$

For example, on the real line $E = \mathbb{R}$ for the Lebesgue measure μ on the unit interval $F = [0,1]$, we have $\alpha = 1$, and (4.3) becomes

$$\mu(B^\delta) \;\ge\; \min\left\{\frac{1}{\delta}\,\mu(B), 1\right\}.$$

For the Cauchy measures μ on $\mathbb{R}^n$ and $\mathbb{R}^\infty$ (cf. examples in section 3.2.2), we have $\alpha = -1$, and then (4.2)-(4.3) with $F = E$ yield

$$\mu(B^\delta) \;\ge\; \frac{\mu(B)}{1 - (1-\delta)(1-\mu(B))}.$$

Note that when E is a Fréchet space and B is Borel, B^δ is universally measurable, so there is no need to use the inner masure.

Let us comment on the extreme values of δ in (4.1) and (4.3). Since the sets B^δ increase for decreasing δ, (4.3) will hold for $\delta = 0$ by continuity, as long as this inequality holds for all $0 < \delta < 1$. In this case, (6.3) with $\delta = 0$ tells as that $\mu(B) > 0 \Rightarrow \mu(B^0) = 1$. Equivalently, after the substitution $A = F \backslash B$ and using Lemma 4.3.2, we get $\mu(A_0) = 0$, that is,

$$\mu\Big\{x \in F : m_\Delta(A) = 1, \text{ for any interval } \Delta \subset F \text{ such that } x \in \Delta\Big\} \;=\; 0,$$

as long as $\mu(A) < 1$. This case is however excluded from the formulation of Theorem 4.0.7 by the assumption $\mu^*(A_\delta) > 0$. Note also that in case $\delta = 1$, (4.1) holds automatically, since then $A_1 = A$.

Thus, both Theorem 4.0.7 and Theorem 4.4.1 do not loose generality by assuming that $0 < \delta < 1$ (and we do this below in this section).

Equivalence of Theorem 4.0.7 and Theorem 4.4.1. It is straightforward for $\alpha \le 0$. This case also includes the values $\mu^*(A_\delta) = 0$ in (4.1), since then $\mu_*(B^\delta) = 1$ for $B = F \setminus A$ and thus both (4.1) and (4.3) are immediate.

Consider the case $0 < \alpha \le 1$. For the implication (4.1) $\Rightarrow$ (4.3), let $p = \mu(B)$, $0 < p < 1$. If $\delta \ge \delta_p = 1 - (1-p)^\alpha$, the formula (4.2) should be applied and then (4.3) becomes

$$\mu_*(B^\delta) \,\ge\, 1 - \left[\frac{(1-\mu(B))^\alpha - (1-\delta)}{\delta}\right]^{1/\alpha}. \tag{4.4}$$

Here the right-hand side tends to 1 as $\delta \downarrow \delta_p$, so necessarily $\mu_*(B^{\delta_p}) = 1$ and hence $\mu_*(B^\delta) = 1$ for all $0 \le \delta < \delta_p$. Thus, without loss of generality, (4.3) may be stated as (4.4) for the range $\delta \ge \delta_p$. If $\mu_*(B^\delta) = 1$ there is nothing to prove. If $\mu_*(B^\delta) < 1$, then $\mu^*(A_\delta) > 0$ for the set $A = F \setminus B$. In that case, (4.1) is exactly the same as (4.4).

For the implication (4.3) $\Rightarrow$ (4.1), assume that $\mu^*(A_\delta) > 0$. Then $\mu_*(B^\delta) < 1$ for $B = F \setminus A$ which implies that $\mu(B) < p_0 = 1 - (1-\delta)^{1/\alpha}$ according to the definition of $R_\delta^{(\alpha)}(\mu(B))$. Moreover, again the formula (4.2) should be applied to rewrite the hypothesis (4.3) in the form (4.4), which can in turn be rewritten as (4.1). ∎

Theorem 4.4.1. Using Theorem 4.0.6, let us show how to reduce the desired statement (4.3) to dimension one. Since the sets B^δ may only become larger, when F is getting larger, one may assume that $F = H_\mu$, i.e., the support of μ. Fix $0 < \delta < 1$.

Step 1: First suppose that B is an open set in F such that the boundary ∂B^δ of B^δ in F has μ-measure zero. Fix an arbitrary $p \in (0,1)$. Using the continuity of the functions $R_\delta^{(\alpha)}$, it is sufficient to show that $\mu(B) > p \Rightarrow \mu(D) \ge R_\delta^{(\alpha)}(p)$, where D is the closure of B^δ. If this were not true, we would have

$$\int (1_B - p)\, d\mu > 0, \qquad \int (R_\delta^{(\alpha)}(p) - 1_D)\, d\mu > 0,$$

which is exactly the condition (1.5) for $u = 1_B - p$ and $v = R_\delta^{(\alpha)}(p) - 1_D$ (where 1_A denotes the indicator function of a set A). These functions are lower-semicontinuous, so we may apply Theorem 1.2: For some one dimensional α-concave probability measure ν supported on an interval $\Delta \subset F$, we have (1.6), i.e.,

$$\nu(B) > p, \qquad \nu(D) < R_\delta^{(\alpha)}(\nu(B)).$$

But

$$\nu(D) \geq \nu\big(B^\delta\big) \geq \nu\big((B \cap \Delta)^\delta\big),$$

where $(B \cap \Delta)^\delta$ is the result of the one dimensional dilation operation applied to $B \cap \Delta$ with respect to Δ. Hence, we obtain $\nu((B \cap \Delta)^\delta) < \nu(B)$ which contradicts the relation (4.3) in dimension one.

Step 2: Here we describe one class of open sets to which the previous step may be applied. Let B be a set of the form $T^{-1}(C) \cap F$, where $T : E \to \mathbb{R}^n$ is a continuous linear map and $C \subset \mathbb{R}^n$ is an open polytope ($n \geq 1$ is arbitrary). Then $(T^{-1}(C))^\delta$ represents an intersection of finitely many open half-spaces (Lemma 5.3). If $\mu(B) > 0$, then, by the zero-one law, the boundaries of these half-spaces have μ-measure zero and hence $\mu(\partial B^\delta) = 0$, as well.

More generally, let $B = T^{-1}(C) \cap F$, where C is a finite union of open polytopes in $\mathbb{R}^n$. Then C^δ is also a finite union of open polytopes. Using Lemma 5.3, we obtain that $\mathrm{clos}(B^\delta) \subset T^{-1}(\mathrm{clos}(C^\delta))$, so $\partial B^\delta \subset T^{-1}(\partial C^\delta)$. Again ∂C^δ is contained in finitely many hyperplanes of $\mathbb{R}^n$ and thus $\mu(\partial B^\delta) = 0$.

Step 3: B is an arbitrary open set in F, assuming that F is a convex, compact set. Denote by $\mathcal{G}$ the collection of all cylindrical sets in F described on the last step. Such sets constitute a base in the original topology on F, since the two coincides by the compactness assumption. Hence $B = \cup G$, where the union is over all $G \in \mathcal{G}$ such that $G \subset B$. Since $\mathcal{G}$ is closed under finite unions, one may apply the Radon property which gives

$$\mu(B) = \sup\{\mu(G) : G \in \mathcal{G},\ G \subset B\}.$$

For any G as above, we have $\mu(G^\delta) \geq R_\delta^{(\alpha)}(\mu(G))$, by the previous steps. Hence, we obtain (6.3) for B, as well.

Step 4: B is an arbitrary Borel set in F. By the strengthened Radon property (4.3), it is sufficient to consider the case of a non-empty compact set B, and we may additionally assume that F is compact.

Any open set in F containing $x \in B$ contains this point together with $B(x) \cap F$, where $B(x) = T_x^{-1}(C(x))$. Here $T_x : E \to \mathbb{R}^n$ is a continuous linear map and $C(x)$ is a Euclidean ball in $\mathbb{R}^n$ (with some n depending on x). Using compactness of B, one can compose its finite covering by the sets of the form

$$G = \big(B(x_1) \cup \dots \cup B(x_N)\big) \cap F, \qquad x_j \in B,$$

with full intersection being B. Let $\{U_i\}_{i \in I}$ be a decreasing net indexed by a semi-ordered directed set I such that each U_i represents the intersection of finitely many sets G as above. The latter guarantees that $\mu(U_i) \downarrow \mu(B)$ along the net.

Now, let $\delta < \delta' < 1$. Given $x \in F$, the property $x \notin B^\delta$ means that $m_{[x,y]}(B) = \inf_{i \in I} m_{[x,y]}(U_i) \leq \delta$ for any $y \in F$. In that case, there is $i \in I$ such that $m_{[x,y]}(U_i) < \delta'$, and hence the increasing sets

$$V_i(x) = \big\{y \in F : m_{[x,y]}(U_i) < \delta'\big\}, \quad i \in I,$$

cover F. By the construction, for each i, the function $\varphi(y) = m_{[x,y]}(U_i)$ is of the type

$$\varphi(y) \,=\, \mathrm{mes}\Big\{t \in (0,1) : \forall\, k \leq l \ \ \exists j \leq N_k \quad (1-t)T_{x_{kj}}(x) + tT_{x_{kj}}(y) \in C(x_{kj})\Big\} \tag{4.5}$$

for some continuous linear maps $T_{x_{kj}} : E \to \mathbb{R}^{n(x_{kj})}$ and some Euclidean balls $C(x_{kj})$ in $\mathbb{R}^{n(x_{kj})}$. As the boundaries of Euclidean balls do not contain non-degenerate intervals, any such function φ must be continuous on F. Therefore, all the sets $V_i(x)$ are open in F, so that by compactness, $V_i(x) = F$ for some $i = i(x)$. Thus, given $x \in F \setminus B^\delta$, we have $m_{[x,y]}(U_{i(x)}) < \delta'$ for any $y \in F$, and hence $F \setminus B^\delta$ is contained in

$$\bigcup_i \big\{x \in F : m_{[x,y]}(U_i) < \delta' \ \text{ for all } y \in F\big\}.$$

It follows that B^δ contains the intersection of the open sets

$$U_i^{\delta'} \ = \big\{x \in F : m_{[x,y]}(U_i) > \delta' \ \text{ for some } y \in F\big\}$$

and thus, by the Radon property,

$$\mu_*(B^\delta) \,\geq\, \mu\big(\cap_i U_i^{\delta'}\big) \,=\, \lim_i \mu\big(U_i^{\delta'}\big).$$

On the other hand, by Step 3, $\mu(U_i^{\delta'}) \geq R_{\delta'}^{(\alpha)}(\mu(U_i))$, and taking the limit along the net we get $\mu_*(B^\delta) \geq R_{\delta'}^{(\alpha)}(\mu(B))$. It remains to let $\delta' \downarrow \delta$ and use the contunuity of $R_\delta^{(\alpha)}$ with respect to δ.

4.5 Large and small deviations

As is known, the dilation-type inequality (4.4) of Theorem 4.0.7 may equivalently be stated on functions (which is often more convenient in applications). Namely, with every Borel measurable function u on E with values in the extended line $[-\infty, \infty]$, one associates its "modulus of regularity"

$$\delta_u(\varepsilon) \;=\; \sup\, \mathrm{mes}\big\{t \in (0,1) : |\, u((1-t)x + ty)\,| \leq \varepsilon\, |u(x)|\big\}, \qquad 0 \leq \varepsilon \leq 1,$$

where the supremum is running over all points $x, y \in E$ such that $u(x)$ is finite.

The behavior of δ_u near zero is used to control the probabilities of large and small deviations of u under hyperbolic measures by involving the parameter α, only (cf. [5, 9, 21]). In particular, there is the following recursive functional inequality, which is stated below, in the setting of an abstract complete locally convex space E.

We assume that μ is an α-concave probability measure on E with $-\infty < \alpha \leq 1$ and that u is a Borel measurable, μ-a.e. finite function on E.

Theorem 4.5.1 *Given* $0 < \lambda < \mathrm{ess\ sup}\,|u|$, *for all* $\varepsilon \in (0,1)$,

$$\mu\{|u| > \lambda\varepsilon\} \;\geq\; \Big[\delta\, \mu\{|u| \geq \lambda\}^{\alpha} + (1-\delta)\Big]^{1/\alpha}, \tag{4.1}$$

where $\delta = \delta_u(\varepsilon)$.

In case $\alpha = 0$, this relation turns into

$$\mu\{|u| > \lambda\varepsilon\} \;\geq\; \big(\mu\{|u| \geq \lambda\}\big)^{\delta}. \tag{4.2}$$

Note that for $\alpha \leq 0$, the assumption $\lambda < \mathrm{ess\,sup}\,|u|$ may be removed.

If μ is supported on a convex closed set F in E, the inequalities (4.1)-(4.2) continue to hold when u is defined on F (rather than on the whole space). In that case, in the definition of δ_u the supremum should be taken over all points $x, y \in F$.

Theorem 4.5.1 Let us recall a simple argument based on Theorem 4.0.7. The latter is applied with $F = E$ to the set

$$A = \{x \in E : \lambda\varepsilon < |u(x)| < \infty\}.$$

By the definiton,

$$\begin{aligned} A_\delta &= \{x \in E : m_{[x,y]}(A) \geq 1 - \delta \ \ \forall y \in E\} \\ &= \big\{x \in E : \mathrm{mes}\{t \in (0,1) : \lambda\varepsilon < |u((1-t)x + ty)| < \infty\} \geq 1 - \delta \ \ \forall y \in E\big\}. \end{aligned}$$

Suppose that $\lambda \leq |u(x)| < \infty$. Then, for any $y \in E$, we have $|u((1-t)x+ty)| \leq \lambda\varepsilon \Rightarrow |u((1-t)x+ty)| \leq \varepsilon|u(x)|$, so that

$$\mathrm{mes}\big\{t \in (0,1) : |u((1-t)x+ty)| \leq \lambda\varepsilon\big\} \ \leq$$
$$\mathrm{mes}\big\{t \in (0,1) : |u((1-t)x+ty)| \leq \varepsilon|u(x)|\big\} \ \leq \ \delta_u(\varepsilon).$$

Hence,

$$\mathrm{mes}\big\{t \in (0,1) : \lambda\varepsilon < |u((1-t)x+ty)| < \infty\big\} \ \geq \ 1 - \delta_u(\varepsilon)$$

which implies that $x \in A_\delta$ with $\delta = \delta_u(\varepsilon)$. This gives the inclusion

$$\{x \in E : \lambda \leq |u(x)| < \infty\} \subset A_\delta$$

and also that $\mu_*(A_\delta) > 0$ (due to the assumption on λ). It remains to apply (4.4). ∎

In the next two corollaries we follow [5], cf. also [21]. Denote by m, a median of $|u|$ under μ, i.e., a real number such that

$$\mu\{|u| > m\} \leq \frac{1}{2}, \quad \mu\{|u| < m\} \leq \frac{1}{2}.$$

Corollary 4.5.2 *Assuming that $m > 0$, for all $r > 1$,*

$$\mu\{|u| \geq mr\} \ \leq \left[1 + \frac{2^{-\alpha} - 1}{\delta_u(\frac{1}{r})}\right]^{1/\alpha}. \tag{4.3}$$

When $\alpha = 0$, the right-hand side is understood as the limit at zero, that is,

$$\mu\{|u| \geq mr\} \ \leq \ 2^{-1/\delta_u(\frac{1}{r})}. \tag{4.4}$$

If $\alpha < 0$, the inequality (4.3) may be simplified as

$$\mu\{|u| \geq mr\} \ \leq \ C_\alpha\, \delta_u(1/r)^{-1/\alpha} \tag{4.5}$$

with constant $C_\alpha = (2^{-\alpha} - 1)^{1/\alpha}$. Note $C_\alpha \to \frac{1}{2}$, as $\alpha \to -\infty$. As is easy to see, we also have a uniform bound, such as, for example, $C_\alpha \leq 1$ in the region $\alpha \leq -1$.

Proof To derive (4.3) in case $\alpha \neq 0$, apply (4.1) with $\lambda = mr$ and $\varepsilon = 1/r$. Then $\mu\{|u| > \lambda\varepsilon\} \leq \frac{1}{2}$, and letting $p = \mu\{|u| \geq \lambda\}$, we get $\frac{1}{2} \geq (\delta p^\alpha + (1-\delta))^{1/\alpha}$. It remains to solve this inequality in terms of p. Note that when $\alpha > 0$, necessarily $\frac{1}{2} \geq (1-\delta)^{1/\alpha}$ or $\frac{2^{-\alpha}-1}{\delta} \geq -1$, so the right-hand side of (4.3) makes sense. By a similar argument, (4.4) follows from (4.2) in the log-concave case.

Remark. An inequality of the form (4.5) can also be obtained by using a transportation argument, cf. [11]. With this argument, a slightly weaker variant of (4.4) is derived in [9].

Now, let us turn to the problem of small deviations.

Corollary 4.5.3 *If $m > 0$, for all $0 < \varepsilon < 1$,*

$$\mu\{|u| \leq m\varepsilon\} \;\leq\; C_\alpha\, \delta_u(\varepsilon) \tag{4.6}$$

with constant $C_\alpha = \frac{2^{-\alpha}-1}{-\alpha}$.

Proof One may assume that $\alpha \neq 0$ and $m = 1$. From (4.1) with $\lambda = 1$, we obtain that $\mu\{|u| \leq \varepsilon\} \leq \varphi(x)$, where $\varphi(x) = 1 - (1+x)^{1/\alpha}$ and $x = (2^{-\alpha}-1)\,\delta_u(\varepsilon)$. Since this function is concave in $x > -1$, we have $\varphi(x) \leq \varphi(0) + \varphi'(0)x = \frac{2^{-\alpha}-1}{-\alpha}\;\delta_u(\varepsilon)$. When $\alpha = 0$, (4.6) holds with $C_0 = \lim_{\alpha\to 0} C_\alpha = \log 2$. ∎

Finally, let us illustrate Corollaries 4.5.2- 4.5.3 on the example of the semi-norms.

Lemma 4.5.4 *If u is a Borel measurable semi-norm on E (not identically zero), then*

$$\delta_u(\varepsilon) = \frac{2\varepsilon}{1+\varepsilon}, \qquad 0 < \varepsilon \leq 1.$$

Proof One may assume that both $u(x)$ and $u(y)$ are finite in the definition of δ_u. Moreover, it is a matter of normalization alone, to assume that $c = u(y) \leq u(x) = 1$. Then, by the triangle inequality,

$$u((1-t)x + ty) \;\geq\; |(1-t)u(x) - tu(y)| \;=\; |(1+c)t - 1|,$$

so

$$\begin{aligned}\mathrm{mes}\big\{t \in (0,1) : u((1-t)x + ty) \leq \varepsilon\, u(x)\big\} &\leq \mathrm{mes}\big\{t \in (0,1) : |(1+c)t - 1| \leq \varepsilon\big\} \\ &= \min\{t_1, 1\} - t_0,\end{aligned}$$

where $t_1 = \frac{1+\varepsilon}{1+c}$, $t_0 = \frac{1-\varepsilon}{1+c}$. In case $c \geq \varepsilon$, we have $t_1 - t_0 = \frac{2\varepsilon}{1+c} \leq \frac{2\varepsilon}{1+\varepsilon}$. In case $c \leq \varepsilon$, similarly $1 - t_0 = \frac{c+\varepsilon}{1+c} \leq \frac{2\varepsilon}{1+\varepsilon}$. Thus, $\delta_u(\varepsilon) \leq \frac{2\varepsilon}{1+\varepsilon}$ in both cases. Here, the equality is attained by taking $y = -x$ with $0 < u(x) < \infty$.

Any Borel measurable semi-norm u on E is generated by a centrally symmetric, Borel measurable, convex set B in E, so that

$$B = \{x \in E : u(x) \leq 1\}.$$

Moreover, u is μ-a.e. finite, if and only if $\mu(B) > 0$ in which case the linear hull of B has μ-measure 1 (by the zero-one law). We are then in position to apply Corollary 4.5.2. More conveniently, starting from (4.1) with $\lambda = r$ and $\varepsilon = \frac{1}{r}$ $(r > 1)$, Lemma 4.5.4 gives

$$\begin{aligned}1 - \mu(B) &= \mu\{u(x) > 1\} \\ &\geq \big[\delta\, \mu\{u(x) \geq r\}^{\alpha} + (1-\delta)\big]^{1/\alpha} \\ &\geq \big[\delta\, (1 - \mu(rB))^{\alpha} + (1-\delta)\big]^{1/\alpha}, \qquad \delta = \frac{2}{r+1}.\end{aligned}$$

At this step, the assumption $\mu(B) > 0$ may be removed. Recalling also Corollary 4.5.3, we arrive at:

Corollary 4.5.5 *Given a symmetric, Borel measurable, convex set B in E, for all $r > 1$,*

$$1 - \mu(B) \,\geq\, \left[\frac{2}{r+1}\, \big(1 - \mu(rB)\big)^{\alpha} + \frac{r-1}{r+1}\right]^{1/\alpha}. \tag{4.7}$$

In the limit case $\alpha = 0$, the above is the same as

$$1 - \mu(rB) \,\leq\, \big(1 - \mu(B)\big)^{(r+1)/2}.$$

This inequality is due to Lovász and Simonovits [31] in case of Euclidean balls B in $\mathbb{R}^n$. Guédon and also found a precise relation in the case $\alpha > 0$. Namely, (4.7) is solved in terms of $1 - \mu(rB)$ as

$$1 - \mu(rB) \,\leq\, \max{}^{1/\alpha}\Big\{\frac{r+1}{2}\,\big(1 - \mu(B)\big)^{\alpha} - \frac{r-1}{2},\, 0\Big\}.$$

As for the range $\alpha < 0$, (7.7) may be then rewritten as

$$1 - \mu(rB) \,\leq\, \Big[\frac{r+1}{2}\,\big(1 - \mu(B)\big)^{\alpha} - \frac{r-1}{2}\Big]^{1/\alpha}.$$

These large deviations bounds provide a sharp form of Borell's Lemma 3.1 in

Let us also mention an immediate consequence from Corollary 4.5.3 and Lemma 4.5.4 concerning measures of small balls.

Corollary 4.5.6 *Given a symmetric, Borel measurable, convex set B in E such that $\mu(B) \leq \frac{1}{2}$, we have*

$$\mu(\varepsilon B) \,\leq\, C_\alpha\, \varepsilon \qquad (0 \leq \varepsilon \leq 1)$$

with constant $C_\alpha = \frac{2(2^{-\alpha}-1)}{-\alpha}$.

Chapter 5

Discrete Brunn-Minkowski

In order develop a notion of curvature for a discrete graph G, we will borrow notions from Riemann Manifolds. For a function $\varphi : \{0, 1, \ldots, n\} \to G$ we will define its length $l = l(\varphi) = \sum_1^n d(\varphi(i), \varphi(i-1))$. In the case that $d(\varphi(i), \varphi(i+1)) = 1$ so that $l(\varphi) = n$, we will call φ a path. Furthermore, we will call the path φ a (distance minimizing) geodesic with initial point $\varphi(0)$ and end point $\varphi(n)$ when ψ a path with $\psi(0) = \varphi(0)$ and $\psi(k) = \varphi(n)$ implies $l(\psi) \geq l(\varphi)$. For $a, b \in G$, we will denote the space of all geodesics with initial point a, and end point b by Γ_{ab}

We will also need a notion of a midpoint. Given $a, b \in G$, we define [1] the set of t-midpoints by

$$\mathcal{M}_t(a, b) = \bigcup_{\gamma \in \Gamma_{ab}} \gamma(\lfloor t/l \rfloor) \cup \gamma(\lceil t/l \rceil). \tag{5.1}$$

More generally for $A, B \subseteq G$,

$$\mathcal{M}_t(A, B) = \bigcup_{(a,b) \in A \times B} \mathcal{M}_t(a, b). \tag{5.2}$$

We can now present our definition of κ curvature lower bound.

Definition A graph G will be said to have a κ-curvature bound when $A, B \subseteq G$ implies

[1] Notice we are using the usual notation for the floor and ceiling of a real number. That is $\lfloor x \rfloor = \max\{z \in \mathbb{Z} : z \leq x\}$, while $\lceil x \rceil = \min\{z \in \mathbb{Z} : z \geq x\}$

that

$$|\mathcal{M}_t(A,B)| \geq |A|^{1-t}|B|^t e^{\kappa t(1-t)\frac{d^2(A,B)}{2}} \tag{5.3}$$

We have defined $\mathcal{M}_t$ to mimic a convex combination of two sets, thus recovering a sense of averaging for the set $\mathcal{P}(G)$. Our definition of κ-curvature bound is in the spirit of $\Phi : \mathcal{P}(G) \to \mathbb{R}$ by $A \mapsto \log|A|$ being a κ-concave in the sense described in the appendix. That is

$$\log|\mathcal{M}_t(A,B)| \geq (1-t)\log|A| + t\log|B| + \kappa t(1-t)d(A,B)/2$$

Noticing that $\log|A|$ is exactly the entropy of uniform measure on A. With this identification of a set with the uniform random variable on said set, it is natural to wonder whether a more abstract approach can be formulated. This brings us to the κ- convexity of entropy, we might ask if $\{\mu_t\}_{t\in[0,1]}$ a geodesic in some sort of Kantorovich space on $\mathcal{P}(G)$ then a κ-curvature holds when the entropy functional is κ-concave. That is,

$$Ent(\mu_t) \geq (1-t)Ent(\mu_0) + tEnt(\mu_1) + \kappa t(1-t)W^2(\mu_0,\mu_1)/2 \tag{5.4}$$

or even more generally that the relative entropy $H(\cdot|\nu)$ with respect to some reference measure ν is κ-convex;

$$H(\mu_t|\nu) \leq (1-t)H(\mu_0|\nu) + tH(\mu_1|\nu) - \kappa t(1-t)W^2(\mu_0,\mu_1)/2. \tag{5.5}$$

5.1 Examples

5.1.1 $\mathbb{Z}, \kappa = 0$

Theorem 5.1.1 *For $A, B \subseteq \mathbb{Z}$*

$$\begin{aligned} \left|\mathcal{M}_t(A,B)\right| \geq & (1-t)|A| + t|B| \qquad (5.6) \\ \geq & |A|^{1-t}|B|^t. \qquad (5.7) \end{aligned}$$

Proof The result on $\mathbb{Z}$ can be borrowed from the result on $\mathbb{R}$, by noticing that for $X \subseteq \mathbb{Z}$ $m(X_{\frac{1}{2}}) = |X|$. Using this observation, and the Brunn-Minkowski on $\mathbb{R}$ we have already

$$\begin{aligned} (1-t)|A| + t|B| &= (1-t)m(A_{\frac{1}{2}}) + tm(B_{\frac{1}{2}}) \qquad (5.8) \\ &\leq m((1-t)A_{\frac{1}{2}} + tB_{\frac{1}{2}}) \qquad (5.9) \end{aligned}$$

Thus, it suffices to prove $(1-t)A_{\frac{1}{2}} + tB_{\frac{1}{2}} = ((1-t)A + tB)_{\frac{1}{2}} \subseteq \mathcal{M}_t(A,B)_{\frac{1}{2}}$. This is obvious when inspected element wise as,

$$((1-t)a + tb)_{\frac{1}{2}} = [(1-t)a + tb - \frac{1}{2}, (1-t)a + tb + \frac{1}{2}] \tag{5.10}$$

while

$$\mathcal{M}_t(a,b)_{\frac{1}{2}} = \left[\lfloor (1-t)a + tb \rfloor - \frac{1}{2}, \lceil (1-t)a + tb \rceil + \frac{1}{2}\right] \tag{5.11}$$

Theorem 5.1.2 *Suppose A, B have the following decompositions into separated intervals $A = \cup_{i=0}^{N} A_i$, $B = \cup_{j=0}^{M} B_j$. Then*

$$|\mathcal{M}_t(A,B)| \geq (1-t)|A| + t|B| + (t \wedge (1-t))\,(N+M)$$

Proof When $N + M = 0$ the result can be computed directly. So we proceed by induction, if $\mathcal{M}_t(A,B)$ is an interval, then

$$\begin{aligned}
|\mathcal{M}_t(A,B)| &= |\mathcal{M}_t(co(A), co(B))| \\
&\geq (1-t)|co(A)| + t|co(B)| \\
&\geq (1-t)|A| + t|B| + (1-t)N + tM.
\end{aligned}$$

If $\mathcal{M}_t(A,B)$ is not an interval, then since $\mathcal{M}_t(A_0, B_0)$ is an interval there exists k, l such that

$$\mathcal{M}_t(\cup_{i=0}^{k} A_i, \cup_{j=0}^{l} B_j)$$

is an interval, but (for definiteness)

$$\mathcal{M}_t(\cup_{i=0}^{k+1} A_i, \cup_{j=0}^{l} B_j)$$

is not. But if $\mathcal{M}_t(A_{k+1}, B_l)$ is not separated from $\mathcal{M}_t(\cup_{i=0}^{k} A_i, \cup_{j=0}^{l} B_j)$ then we have contradiction, as the union of the two unseperated intervals will be an interval, and this would imply $\mathcal{M}_t(\cup_{i=0}^{k+1} A_i, \cup_{j=0}^{l} B_j)$ is an interval. So we have

$$\mathcal{M}_t(\cup_{i=0}^{k} A_i, \cup_{j=0}^{l} B_j) < \mathcal{M}_t(A_{k+1}, B_l) \leq \mathcal{M}_t(\cup_{k+1}^{N} A_i, \cup_{l}^{M} B_j)$$

Thus,

$$
\begin{aligned}
|\mathcal{M}_t(A,B)| \geq & |\mathcal{M}_t(\cup_{i=0}^k A_i, \cup_{j=0}^l B_j)| + |\mathcal{M}_t(\cup_{k+1}^N A_i, \cup_l^M B_j)| \\
\geq & (1-t)|\cup_{i=0}^k A_i| + t|\cup_{j=0}^l B_j| + t \wedge (1-t)(k+l) \\
& + (1-t)|\cup_{k+1}^N A_i| + t|\cup_l^M B_j| + t \wedge (1-t)(N-k-1+M-l) \\
= & (1-t)|A| + t|B| + t \wedge (1-t)(N+M-1) + t|B_l| \\
\geq & (1-t)|A| + t|B| + t \wedge (1-t)(N+M)
\end{aligned}
$$

We will now show that the 1-dimensional Brunn-Minkowski can be derived for $\mathbb{Z}/n$ from the 1-dimensional result on $\mathbb{Z}$.

Theorem 5.1.3 *For nonempty $A, B \subseteq \mathbb{Z}/n$*

$$|\mathcal{M}_t(A,B)| \geq (1-t)|A| + t|B|.$$

Proof Notice the isomorphism between $\mathbb{Z}/n$ and the discrete subset of S^1, $\{e^{\frac{2\pi ik}{n}}\}$, can be made an isometry by multiplying the usual metric on $\mathbb{S}^1$ by $\frac{n}{2\pi}$, this allows us to identify points of $\mathbb{Z}/n$ by their angle in $\mathbb{S}^1$.

First a special case, if one of the sets A or B, consists of a single element x, cut S^1 at the antipodal point to x, and consider as A and B as subsets of $\mathbb{Z}$. This will only reduce midpoints, and the result follows.

With this simple case removed, let us consider two other tractable situations.

Convex Subsets of $\mathbb{Z}/n$ Can Be Considered as Subsets of $\mathbb{Z}$

Should both A and B be subsets of a halfspace, $[y, y+\pi] = [e^{i\theta}, e^{i(\theta+\pi)}]$, then the points of $\mathcal{M}_t(A,B) \cap [y, y+\pi]$ coincide with the midpoints of A, B considered as subsets of the interval in $\mathbb{Z}$; $[n\theta/2\pi, n(\theta+\pi)/2\pi]$ which we denote as $A_{\mathbb{Z}}$ and $B_{\mathbb{Z}}$ respectively. Now, computing;

$$
\begin{aligned}
|\mathcal{M}_t(A,B)| &\geq |\mathcal{M}_t(A,B) \cap [y, y+\pi]| && (5.12) \\
&= |\mathcal{M}_t(A_{\mathbb{Z}}, B_{\mathbb{Z}})| && (5.13) \\
&\geq (1-t)|A| + t|B|. && (5.14)
\end{aligned}
$$

When A and B can be Simultaneously Split

Given $X \subseteq \mathbb{Z}/n$, and C_0, C_1 convex subsets of $\mathbb{Z}/n$, such that $C_0 \cup C_1 = \mathbb{Z}/n$ and $C_0 \cap C_1 = \emptyset$, we will call the sets a *convex splitting of* X if $X_i \coloneqq X \cap C_i \neq \emptyset$. When $X = \mathbb{Z}/n$ or the subset is to be understood or not of importance, we will simply call the pair of C_i's a *convex splitting*. In the case that the same pair (C_0, C_1) form a convex splitting of both A and B, we will refer to the pair as a *simultaneous splitting*.

Now suppose that for (A, B), simultaneous splitting (C_0, C_1) exists. Then, making use of the result above for convex subsets,

$$
\begin{aligned}
|\mathcal{M}_t(A, B)| &= |\mathcal{M}_t(A_0 \cup A_1, B_0 \cup B_1)| && (5.15)\\
&\geq |\mathcal{M}_t(A_0, B_0) \cap C_0| + |\mathcal{M}_t(A_1, B_1) \cap C_1| && (5.16)\\
&\geq (1-t)|A_0| + t|B_0| + (1-t)|A_0| + t|B_0| && (5.17)\\
&= (1-t)|A| + t|B|. && (5.18)
\end{aligned}
$$

"The General Case"

What we will actually show here, is that for $A, B \subseteq \mathbb{Z}$, there is no general case; that is with $|A|, |B| \geq 2$ there is either a containing half plane, or a simultaneous splitting.

To start if every convex splitting of $\mathbb{Z}/n$ is a splitting of A, then by choosing any splitting of B, we will arrive at a simultaneous splitting, and the result follows from case 3. Thus, there exists some convex splitting of $\mathbb{Z}/n$ that does not split A, and hence we may assume A is contained in a halfspace. In order to choose convenient coordinates assume $0 \in A$ and that $A \subseteq [0, \pi]$.

Notice that $(0, \pi]$,$(\pi, 0]$ is a convex splitting of A, should it also split B we are done, so we may assume that B is contained in either $(0, \pi]$ or $(\pi, 0]$. If $B \subseteq (0, \pi]$ then $A \cup B \subseteq [0, \pi]$ and case 2 applies. So it must be the case that $B \subseteq (\pi, 0]$.

Take $\alpha = \max_{(0,\pi]} A$ then $[\alpha, \alpha+\pi)$,$[\alpha+\pi, \alpha)$ also represents a convex splitting of A, from which it follows that $B \subseteq [\alpha, \alpha+\pi) \cap (\pi, 0] = (\pi, \alpha+\pi)$. Letting $b = \min_(\pi, \alpha+\pi)B$, then via its definition $(b-\pi, b]$, $(b, b+\pi]$ is a convex splitting of B. But, since $\alpha \in (b-\pi, b]$ and $0 \in (b, b+\pi]$ this is a splitting of A. Hence a simultaneous splitting exists for (A, B) and case 3 applies.

5.1.2 Prékopa-Leindler for G satisfying a 1-Dimensional Brunn-Minkowski

We will say that a graph G satisfies a Prékopa-Leindler inequality (PLI) if given $f, g, h : G \to \mathbb{R}^+$ such that for $z \in \mathcal{M}_t(x, y)$,

$$f(z) \geq g^{1-t}(x)h^t(y), \tag{5.19}$$

Then

$$\int f \geq (\int g)^{1-t}(\int h)^t.$$

Where we use the notation $\int \varphi := \sum_{x \in G} \varphi(x)$.

Theorem: *If $A, B \subseteq G$ implies $|\mathcal{M}_t(A, B)| \geq (1-t)|A| + t|B|$, then G satisfies PLI.*

Proof:

The proof is the same as for $\mathbb{R}$. By homogeneity, it suffices to consider f, g, h satisfying the hypothesis such that $\max g = \max h = 1$. Choose $z \in \mathcal{M}_t(\{g > \lambda\}, \{h > \lambda\})$, then$z \in \mathcal{M}_t(x, y)$ for some x, y such that $g(x) > \lambda, h(y) > \lambda$. Thus, by the hypothesis on f, g, h, $f(z) \geq g(x)^{1-t}h(y)^t > \lambda$. Hence

$$\{f > \lambda\} \supseteq \mathcal{M}_t(\{g > \lambda\}, \{h > \lambda\})$$

Now by Fubini-Tonelli, our set theoretic inclusion, the Brunn-Minkowski hypothesis, and AM-GM inequality,

$$\int f = \int_0^\infty |\{f > \lambda\}| d\lambda \tag{5.20}$$

$$\geq \int_0^1 |\{f > \lambda\}| d\lambda \tag{5.21}$$

$$\geq \int_0^1 |\mathcal{M}_t(\{g > \lambda\}, \{h > \lambda\})| d\lambda \tag{5.22}$$

$$\geq (1-t) \int_0^1 |\{g > \lambda\}| d\lambda + t \int_0^1 |\{h > \lambda\}| d\lambda \tag{5.23}$$

$$\geq \left(\int g \right)^{1-t} \left(\int h \right)^t . \tag{5.24}$$

In particular both $\mathbb{Z}$ and $\mathbb{Z}/n$ satisfy PLI.

5.2 Extensions

In the continuous case, one of the more useful properties of the Prékopa-Leindler inequality is that it tensorizes. In the discrete setting, such behavior is less clear. In the continuous case, where exact midpoints are available

$$\mathcal{M}_t(A_1, B_1) \times \mathcal{M}_t(A_2, B_2) \subseteq \mathcal{M}_t(A_1 \times A_2, B_1 \times B_2)$$

In the discrete setting the error induced by approximate midpoints proliferates, and the above inequality is not true in general. For example take the product two point space $\{0, 1\}$, $A_1 = A_2 = \{0\}$ and $B_1 = B_2 = \{1\}$. By our definitions $\mathcal{M}_{\frac{1}{2}}(A_1, B_1) = \mathcal{M}_{\frac{1}{2}}(A_2, B_2) = \{0, 1\}$ so that

$$\mathcal{M}_{\frac{1}{2}}(A_1, B_1) \times \mathcal{M}_{\frac{1}{2}}(A_2, B_2) = \{0, 1\}^2.$$

But

$$\mathcal{M}_{\frac{1}{2}}(A_1 \times A_2, B_1 \times B_2) = \mathcal{M}_{\frac{1}{2}}(\{0, 0\}, \{1, 1\}) = \{1, 0\} \cup \{0, 1\}.$$

Worth mentioning here is that the set above gives an example of two convex sets $\{0, 0\}$ and $\{1, 1\}$ whose midpoint set is not convex.

A further peculiarity, is the fact that the dimensional Brunn-Minkowski fails for $\mathbb{Z}^2$. That is there exists $A, B \subseteq \mathbb{Z}^2$ and $t \in (0,1)$ such that

$$|\mathcal{M}_t(A,B)| < ((1-t)|A|^{\frac{1}{2}} + t|B|^{\frac{1}{2}})^2$$

In particular, take $t = \frac{1}{2}$, $A = \{(0,0)\}$, and $B = \{x : |x_1| + |x_2| \leq 2$, so that $|\mathcal{M}_t(A,B)| = \{x : |x_1| + |x_2| \leq 1\}$. Its easily computed that

$$|\mathcal{M}_t(A,B)| = 5 < \frac{(1+13^{\frac{1}{2}})^2}{4} = ((1-t)|A|^{\frac{1}{2}} + t|B|^{\frac{1}{2}})^2$$

References

[1] L.M. Arutyunyan and E.D. Kosov. Estimates of integral norms of polynomials on spaces with convex measures. *to appear Sbornik Math.*, 2015.

[2] L.M. Arutyunyan and E.D. Kosov. On polynomials on spaces with convex measures. *to appear Doklady Math.*, 2015.

[3] S. G. Bobkov and J. Melbourne. Localization for infinite dimensional hyperbolic measures (russian). *Doklady Akademii Nauk*, 462(3):261–263, 2015. English translation in *Doklady Mathematics*, 91(3), 2015.

[4] S. G. Bobkov and J. Melbourne. Hyperbolic measures on infinite dimensional spaces. *Probability Surveys*, Accepted upon minor revision, arXiv:1405.2961.

[5] S. G. Bobkov and F. L. Nazarov. Sharp dilation-type inequalities with a fixed parameter of convexity. *Journal of Mathematical Sciences*, 152(6):826–839, 2008.

[6] Sergey G. Bobkov. Isoperimetric problem on the real line. *Zapiski Nauchnykh Seminarov POMI*, 216:5–9, 1994.

[7] Sergey G. Bobkov. Remarks on the growth of l p-norms of polynomials. In *Geometric aspects of functional analysis*, pages 27–35. Springer, 2000.

[8] Sergey G Bobkov. Some generalizations of prokhorov's results on khinchin-type inequalities for polynomials. *Theory of Probability & Its Applications*, 45(4):644–647, 2001.

[9] Sergey G Bobkov. Large deviations via transference plans. *Adv. Math. Res*, 2:151–175, 2003.

[10] Sergey G. Bobkov. Localization proof of the bakry–ledoux isoperimetric inequality and some applications. *Theory of Probability & Its Applications*, 47(2):308–314, 2003.

[11] Sergey G Bobkov. Large deviations and isoperimetry over convex probability measures with heavy tails. *Electron. J. Probab*, 12:1072–1100, 2007.

[12] Sergey G. Bobkov. On isoperimetric constants for log-concave probability distributions. In *Geometric aspects of functional analysis*, pages 81–88. Springer, 2007.

[13] Sergey G. Bobkov and Mokshay Madiman. Concentration of the information in data with log-concave distributions. *The Annals of Probability*, 39(4):1528–1543, 2011.

[14] V.I. Bogachev, O.G. Smolyanov, and V.I. Sobolev. Topological vector spaces and their applications (russian). *Moscow, Izhevsk*, 2012.

[15] Vladimir I Bogachev. *Measure theory*, volume 1. Springer Science & Business Media, 2007.

[16] Christer Borell. Convex measures on locally convex spaces. *Arkiv för Matematik*, 12(1):239–252, 1974.

[17] Christer Borell. Convex set functions ind-space. *Periodica Mathematica Hungarica*, 6(2):111–136, 1975.

[18] Herm Jan Brascamp and Elliott H Lieb. On extensions of the brunn-minkowski and prékopa-leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *Journal of Functional Analysis*, 22(4):366–389, 1976.

[19] Yuriui Dmitrievich Burago, Viktor A. Zalgaller, and A.B. Sossinsky. *Geometric inequalities*, volume 1988. Springer Berlin, 1988.

[20] Luis A Caffarelli. Monotonicity properties of optimal transportation and the fkg and related inequalities. *Communications in Mathematical Physics*, 214(3):547–563, 2000.

[21] Matthieu Fradelizi. Concentration inequalities for s-concave measures of dilations of borel sets and applications. *Electron. J. Probab*, 14:2068–2090, 2009.

[22] Matthieu Fradelizi and Olivier Guédon. The extreme points of subsets of s-concave probabilities and a geometric localization theorem. *Discrete & Computational Geometry*, 31(2):327–335, 2004.

[23] N. Gozlan, J. Melbourne, W. Perkins, C. Roberto, P. Samson, and P. Tetali. Discrete ricci curvature and functional inequalities on graphs: a survey. *in preperation*.

[24] N. Gozlan, J. Melbourne, W. Perkins, C. Roberto, P. Samson, and P. Tetali. Some versions of the prékopa leindler inequality on $\mathbb{Z}$. *In preperation*, preprint `math.umn.edu/~melbo013`.

[25] Nathael Gozlan, Cyril Roberto, Paul-Marie Samson, and Prasad Tetali. Displacement convexity of entropy and related inequalities on graphs. *Probability Theory and Related Fields*, pages 1–48, 2012.

[26] M Gromov and VD Milman. Generalization of the spherical isoperimetric inequality to uniformly convex banach spaces. *Compositio mathematica*, 62(3):263–282, 1987.

[27] Olivier Guédon. Kahane-khinchine type inequalities for negative exponent. *Mathematika-London*, 46(1 ISSU 91):165–174, 1999.

[28] Hugo Hadwiger and D. Ohmann. Brunn-minkowskischer satz und isoperimetrie. *Mathematische Zeitschrift*, 66(1):1–8, 1956.

[29] R. Kannan, L. Lovász, and M. Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete & Computational Geometry*, 13(1):541–559, 1995.

[30] L.V. Kantorovich and G.P. Akilov. Functional analysis. 1982.

[31] L. Lovász and M. Simonovits. Random walks in a convex body and an improved volume algorithm. *Random structures & algorithms*, 4(4):359–412, 1993.

[32] László Lovász and Miklós Simonovits. Random walks in a convex body and an improved volume algorithm. *Random structures & algorithms*, 4(4):359–412, 1993.

[33] Paul-André Meyer. *Probability and potentials.* Blaisdell Publishing Company Waltham (Mass.), 1966.

[34] Yann Ollivier and Cédric Villani. A curved brunn–minkowski inequality on the discrete hypercube, or: What is the ricci curvature of the discrete hypercube? *Siam Journal on Discrete Mathematics*, 26(3):983–996, 2012.

[35] Lawrence E Payne and Hans F. Weinberger. An optimal poincaré inequality for convex domains. *Archive for Rational Mechanics and Analysis*, 5(1):286–292, 1960.

[36] Robert R Phelps. *Lectures on Choquet's theorem*, volume 1757. Springer Science & Business Media, 2001.

[37] Walter Rudin. Functional analysis. international series in pure and applied mathematics, 1991.

[38] M. Sodin, F. Nazarov, and A. Volberg. The geometric kannan-lovász-simonovits lemma, dimension-free estimates for the distribution function of the values of polynomials, and the distribution of the zeros of random analytic functions. *St. Petersburg Math. J*, 14:351–366, 2003.

[39] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer, 2008.

Appendix A

Prerequisite

Definition 1 *We will call $A \subset E$ an affine subspace when $A = w + V$ for some $w \in E$ and V a vector subspace of E.*

Theorem A.0.1 *The following are equivalent;*

1. *A is an affine subspace*
2. *$\lambda A + (1-\lambda)A \subset A$ for $\lambda \in \mathbb{R}$*
3. *$\lambda_1 A + \dots \lambda_n A \subset A$ for $\lambda_1 + \cdots + \lambda_n = 1$*

Proof ($1 \Rightarrow 2$) Suppose A is an affine subspace, with $A = w + V$ for $v \in E$ and V a vector subspace. Given $a_0, a_1 \in A$, $\lambda \in \mathbb{R}$, then $a_0 = v_0 + w, a_1 = v_1 + w$ for some $w_0, w_1 \in V$. Then,

$$\begin{aligned}(1-\lambda)a_0 + \lambda a_1 =& (1-\lambda)(w_0 + v) + \lambda(w_1 + v)\\ =& (1-\lambda)w_0 + \lambda w_1 + v = w_2 + v\end{aligned}$$

Where $w_2 = (1-\lambda)w_0 + \lambda w_1$ is an element of the vector space V.

($2 \Rightarrow 3$) By induction, suppose that $\lambda_1 + \cdots + \lambda_n = 1$ and that the result hold for any smaller collection. Without loss of generality assume that $\lambda = \lambda_1 + \dots \lambda_{n-1} \neq 0$, and take $(a_i)_i \in A^n$. By the induction hypothesis $a = \frac{\lambda_1}{\lambda} a_1 + \cdots + \frac{\lambda_{n-1}}{\lambda} a_{n-1}$ is an element of A. Hence,

$$\begin{aligned}\lambda_1 a_1 + \cdots + \lambda_{n-1}a_{n-1} + \lambda_n a_n =& \lambda\left(\frac{\lambda_1}{\lambda}a_1 + \cdots + \frac{\lambda_{n-1}}{\lambda}a_{n-1}\right) + \lambda_n a_n \\ =&(1-\lambda_n)a + \lambda_n a_n \in A\end{aligned}$$

($3 \Rightarrow 1$) Fix $a \in A$. We claim that $V_a = A - a$ is a vector space. For an element $x \in V_a$ and a scalar λ

$$\lambda x = \lambda(a_0 - a) = \lambda a_0 + (1-\lambda)a - a \tag{A.1}$$

But $\lambda a_0 + (1-\lambda)a \subset \lambda A + (1-\lambda)A \subset A$. So V_a is closed under scalar multiplication. To show it is closed under vector addition compute for $v, w \in V_a$

$$v + w = (a_v - a) + (a_w - a) \tag{A.2}$$

$$= (a_v + a_w + (-1)a) - a \tag{A.3}$$

$a_v + a_w + (-1)a$ is an element of A since the coefficients $1 + 1 - 1 = 1$.

Corollary A.0.2 *If $V_{a'} = A - a'$ and $V_a = A - a$ for an affine subspace A, then $V_{a'} = V_a$.*

Proof Given $x \in V_a$, $x = a_x - a$ for some $a_x \in A$ and hence $a_x - a + a' \in A$ by the characterization of affine subspaces in terms of linear combinations. Thus $x \in V_{a'}$. Hence $V_a \subset V_{a'}$ and by symmetry $V_a = V_{a'}$.

The previous corollary insures that the following is well defined.

Definition For an affine subspace $A = V + a$ define the dimension of A to be the dimension of its related vector space $dim(A) = dim(V)$.

Definition For $X \subset E$, define the **affine hull**, $aff(X)$ to be the smallest affine subspace containing X. More precisely, $aff(X)$ is the intersection of all affine subspaces containing X.

It is not difficult to see that

$$aff(X) = \left\{ v : v = \sum_{i=1}^{n} \lambda_i x_i, \lambda_1 + \cdots + \lambda_n = 1, x_i \in X \right\}. \tag{A.4}$$

Finally we can define the dimension of a measure μ.

Definition Define the **dimension** of μ, denoted $dim(\mu)$ as $dim(aff(supp(\mu)))$.

Theorem A.0.3 *Suppose that μ is a probability measure, then for f positive and measurable, then the function $F(\alpha) = \left(\int f^{\alpha} d\mu\right)^{\frac{1}{\alpha}}$ is non-decreasing.*

Proof Starting $0 < \alpha < \beta$, so that $x \mapsto x^{\frac{\beta}{\alpha}}$ is convex. Applying Jensen's inequality,

$$\int f^{\beta} d\mu = \left(\int f^{\alpha}\right)^{\frac{\beta}{\alpha}} \tag{A.5}$$

$$\geq \left(\int f^{\alpha} d\mu\right)^{\frac{\beta}{\alpha}}. \tag{A.6}$$

So that $F(\beta) = (\int f^{\beta} d\mu)\overset{1\beta}{\geq}(\int f^{\alpha} d\mu)\overset{1\alpha}{=}F(\alpha)$.

A.1 Information Theory

For a discrete random variable X taking n different values, the ith event occuring with probability p_i. Define the Entropy of X;

$$Ent(X) = \sum_i p_i \log(\frac{1}{p_i})$$

A motivation for such a definition is the following. X is a random variable whose outcome we would like to communicate to a second party. To do so we will encode each result with a string of zeros and ones to be transmitted and decoded. $\log(\frac{1}{p_i})$ represents the optimal length of the string associated to event i. Notice that $\log(\frac{1}{x})$ is decreasing, so that we use short strings on likely events, at the expense of using relatively long strings on unlikely events. The entropy of a random variable represents the average length of the message sent. This quantifies a sort of "cost" to transmit the result of a random variable, and thus the value of knowing the result of the random variable. In that sense the entropy is the amount of information "in" a random variable.

The relative entropy, $H(\mu|\nu)$ for $\mu << \nu$ describes the average waste in coding as though a phenomena has distribution ν rather than its true distribution μ. Using

$p_i = \mu(x_i)$ and $q_i = \nu_i$ then

$$H(\mu|\nu) = \sum_i p_i \log(\frac{1}{q_i}) - \sum_i p_i \log(\frac{1}{p_i}) \tag{A.7}$$

$$= \sum_i p_i/q_i \log(p_i/q_i)q_i \tag{A.8}$$