Sudan University of Science and Technology

College of Graduate Studies

Rough Curvature Bounds and Free Functional inequalities

قيود الانحناء الخشن ومتباينات الدوال الحرة

A thesis Submitted in Fulfillment of the Requirements for the degree of Philosophy in Mathematics

By

Mohammed Ali Abdlkariem Arabi

Supervisor

Prof.Dr: Shawgy Hussein Abd Alla

March 2014

Dedication

To my glory is to my Mother and Father.

Acknowledgement

I wish to acknowledge valuable support of my supervisor Prof. Dr. Shawgy Hussein Abd Alla for his guidance and help during writing the dissertation.

My thanks are also due to the staff of Mathematical Department, College of science, Sudan University of Sciences and Technology.

I would like to extend my thanks to the Dean of the College of Sciences, Sudan University of Sciences and Technology for this helpful advice & encouragement throughout this research.

Finally I would like to thank the Bahri University and for the scholarship.

My thanks are due to Babiker Mohammed Elzain at Sudan institute for Natural Sciences.

Abstract

We study the transport inequalities, gradient estimates entropy, and Ricci curvature. A free probability method of the Wasserstein metric on the trace-states space is considered. We give a free Brunn-Minkowski inequality, and show the Talagrand inequality for the semicircular law and energy of eigenvalues of Beta ensembles. We also show the Ricci curvature for metric measure spaces by optimal transport, and consider the mass transportation and rough curvature bounds for discrete spaces. We investigate the combinatorial dimension and certain norms in the method of harmonic analysis, and characterize the relationships between combinatorial measurements and Orlicz norms. Also Characterization of dimension free concentration in terms of transportation and Poincar'e inequalities with dimension free concentration of measure are shown, mass transportation evident of free functional inequalities and free Poincar'e inequalities are confirm.

الخلاصة

درسنا متباينات التنقل ودرجة القصورالحراري لتقديرات الميل وا إنحناء ريكاي. إعتبرنا طريقة الإحتمال الحر لمترك واسرشتاين على فضاء أثر– الحالات. تم إعطاء متباينة برن– منكووسكاي الحرة وتوضيح متباينة تالاجراند لأجل القانون نصف الدائري وفعالية القيم الذاتية لفرق بيتا. أيضاً أوضحنا إنحناء ريكاي لفضاءات القياس المترية بواسطة التنقل الأمثل وا إعتبرنا تنقل الكتلة وحدوديات الإنحناء الخشن للفضاءات المتقطعة. إستقصينا البعد الإندماجي ونظائم معينة في طريقة التحليل التوافقي وشخصنا العلاقات بين القياسات الإندماجية ونظائم أورلش. أوضحنا أيضاً تشخيص التمركز حر البعد بدلالات التنقل ومتباينات بونكارية مع التمركز حر البعد للقياس. تم تأكيد وضوح تنقل الكتلة للمتباينات الدالية الحرة ومتباينات بونكارية الحرة.

Introduction

We present various characterizations of uniform lower bounds for the Ricci curvature of smooth Riemannian manifold in convexity properties of the entropy.

We define free probability analogues of the Wasserstein metric, which extends the classical one. We present one dimensional various of the functional form of the geometric Brunn-Minkowski inequality in free (non-commutative) probability theory. The proof relies on matrix approximation as used recently by Biane and Hiai et al to establish free analogues of the Logarithmic Sobolev and transportation-cost inequalities for strictly convex potentials that are recovered here from the Brunn- Minkowski inequality as in the classical case. We give a short proof of an extension of the free Talagrand transportation –cost inequality to the semicircular which was originally proved [198].The proof is based on a convexity argument and is the spirit of the original Talagrand's approach for the classical counterpart from [179].

We define a notion of measured length space having nonnegative Ricci curvature or having ∞-Ricci curvature bounded below by a real number .We introduce and study rough (approximate) lower curvature bounds for discrete spaces and graphs. This notion agrees with the one introduce in the sense that the metric measure space which is approximated by a sequence of discrete spaces with rough curvature greater than or equals a real constant will have curvature greater than or equals other a real constant.

We study a parameter called combinatorial dimension where appropriate constructions in a harmonic, analytic framework filled "combinatorial " and " analytic " gaps are open between Cartesian products of spectral sets. We establish in a setting of harmonic analysis precise relationships between combinatorial measurement and Orlicz norms.

The aim is to show that a probability measure on R^d concentrate independently of the dimension like a Gaussian measure , if and only if it verify Talagrand's T_2 transportation –cost inequalities .We consider Poincar'e inequalities for non Euclidean matrics on R^d .These inequalities rate between type exponential and Gaussian and beyond. This work is devoted to a direct mass transportation proofs of families of functional inequalities in the context of one dimensional free probability, avoiding random matrix approximation. The inequalities include the free form of the transportation log Sobolev , *HWI* interpolation and Brunn-Minkowski inequalities for strictly convex potentials. Sharp constants and some extended version are put forward.

The contents

Subject	**Page**
Dedication	I
Acknowledgement	II
Abstract	III
Abstract (Arabic)	IV
Introduction	V
The contents	VI
Chapter 1: Transportation Inequalities and Gradient Estimates with Entropy	1
Chapter 2: Free Probability of Wasserstein metric on Trace-State Space	19
Chapter 3: Free Brimn-Minkowski and Talagrand Inequalities	32
Section (3.1): One dimensional Brunn-Minkowski Inequality	32
Section (3.2): Semicircular Law and Energy of the eigenvalues of Beta Ensembles	36
Chapter 4: Ricci Curvature for Metric Measure Spaces	48
Section (4.1): Geometry and Functionals of Wasserstein Space	48
Section (4.2): Ricci Curvature for Measured Length Spaces And Riemannian Manifolds	70
Section (4.3): Mass transportation and rough curvature bounds	99
Chapter 5: Dimension and Relation Between Combinatorial Measurement	124
Section (5.1): Dimension and Norms in Harmonic Analysis	124
Section (5.2) : Combinatorial Measurements and Orlicz Norms	152
Chapter 6: Mass Transportation of free Functional Inequalities and Poincare Inequalities	196
Section (6.1): Characterization of Dimension Free Concentration In Transportation Inequality	196
Section (6.2): Poincare Inequalities and Dimension of Free Concentration of Measure	227
Section (6.3): Mass transportation of Free Functional Inequalities and free Poincare inequalities	276
List of Symbols	330
References	331

Chapter 1

Transport Inequalities and Gradient Estimates with Entropy

The entropy will be considered as a function on the space of probability measures on the Riemannian manifold as well as in terms of transportation inequalities for volume measures, heat kernels, and Brownian motions and in terms of gradient estimates for the heat semigroup**.**

For metric measure spaces there is neither a notion of Ricci curvature nor a common notion of bounds for the Ricci curvature See [104,156,143,251,285,118] (complete, for instance, to Alexandrov's notion of bounds for the sectional curvature for metric spaces [29.166, 4,129,163,145].

We present various characterizations of uniform lower bounds for the Ricci curvature of a smooth Riemannian manifold M in terms of convexity properties of the entropy (considered as a function on the space of probability measures on M) as well as in terms of transportation inequalities for volume measures, heat kernels, and Brownian motions and in terms of gradient estimates for the heat semi group.

In what follows, $(\mathrm{M,g})$ is always assumed to be a smooth, connected, complete Riemannian manifold with dimension n, Riemannian distance $\mathrm{d}(\mathrm{x},\mathrm{y})$, and Riemannian volume $m(dx)=vol(dx)$, for $r\in[1,\infty)$ the L^r-Wasserstein distance of two measures μ_1 and μ_2 on M is defined as:

$$\mathrm{W}_r\left(\mu_1,\mu_2\right):=\inf\left\{\int_{M\times M} d\left(x_1,x_2\right)^r \mu\left(dx_1dx_2\right):\mu\in C\left(\mu_1,\mu_2\right)\right\}^{\frac{1}{r}}$$

Where $C(\mu_1,\mu_2)$ denotes the set of all coupling of μ_1 and μ_2, that is, the set of all measures μ on $M\times M$ with $\mu(A\times M)=\mu_1(A)$ and $\mu(M\times A)=\mu_2(A)$ for all measurable $A\subset M$: see [23].

Here and in what follows, the "measure on M" always means the measure on M equipped with its Borel σ-field. $\mathrm{P}^r(M)$ will denotes the set of probability measures μ on M with $\int_M d(x,y)\mu(dy)<\infty$ for some (hence all) $x\in M$. Equipped with the metric W_r the space $\mathrm{P}^r(M)$ is a geodesic space.

The relative entropy is defined as a function on $\mathrm{P}^r(M)$ by:

$$H(\mathrm{V}) := \int_M \frac{dv}{dx}\log\frac{d\mathrm{V}}{dx}\mathrm{V}ol(dx)$$

If V is absolutely continuous with respect to vol with $\int_M \frac{d\mathrm{V}}{dx}\left[\log\frac{d\mathrm{V}}{dx}\right] + \mathbf{v}ol(dx) < \infty$ and $H(\mathrm{V}) := +\infty$ otherwise. Given an arbitrary geodesic space (X, ρ), a number $K \in R$, and a function $U : X \to (-\infty, +\infty)$, we say that U is K convex if and only if for each (constant speed, as usual) geodesic $\mathrm{V} : [0,1] \to X$ with $U(\mathrm{V}_0) < \infty$ and $U(\mathrm{V}_1) < \infty$ for each $t \in [0,1]$,

$$U(\mathrm{V}_t) \le (1-t)U(\mathrm{V}_0) + tU(\mathrm{V}_1) - \frac{K}{2}t(1-t)\rho^2(\mathrm{V}_0, \mathrm{V}_1).$$

K - Convex function on $\mathrm{P}^2(M)$ are also called displacement K - convex (to make sure that $t \mapsto \mathrm{V}_t$ is really the geodesic with respect to W_r and not the linear interpolation $t \mapsto (1-t)v_0 + t\mathrm{V}_1$ in the space $\mathrm{P}^2(M)$).

Here and henceforth, $p_t(x,y)$ always denotes the heat kernel on M, i.e. the minimal positive fundamental solution to the heat equation $\left(\Delta - \frac{\partial}{\partial t}\right)p_t(x,y) = 0$. It is smooth in (t,x,y) and symmetric in (x,y). And it satisfies $\int_M P_t(x,y)\mathrm{V}d(dx) \le 1$. Hence, it defines a sub probability measure $P_t(x,dy) = P_t(x,y)vol(dy)$ as well as operators $P_t : C_t^\infty(M) \to C^\infty(M)$ and $P_t : L^2(M) \to L^2(M)$. Which are all denoted by the same symbol given $\mu \in \mathrm{P}^r(M)$ and $t > 0$, we define a new measure $\mu_{P_t} \in \mathrm{P}^r(M)$ by:

$$\mu P_t(A) = \int_A \int_M P_t(x,y)\mu(dx)vol(dy).$$

Brownian motion on M is by definition the Markov process with generator $\frac{1}{2}\Delta$. Thus its transition (sub) probabilities are given by $P_{\frac{1}{2}}$.

If the Ricci curvature of the underlying manifold M is bounded from below, then all the $P_t(x,\cdot)$ are probability measures. If the later holds true, we say that the heat kernel and the associated motion has an infinite lifetime.

One obtains contraction in W_r for each $r \in [1,\infty]$ and for any initial data and One obtains path wise contraction for Brownian trajectories.

The advantage of this characterization of Ricci curvature is that it depends only on the basic, robust data: measure and metric. It does not require any heat kernel, any Laplacian, or any Brownian motion. It might be used as a guideline in much more general situations.

For instance, Let (M,d) be an arbitrary separable metric space equipped with a measure m on its Borel σ-field and assume that (2) holds true (with some numbers $K \in \mathbf{R}$ and $n > 0$). Define an operator m_r acting on bounded measurable functions by:

$$m_r f(x) = \int_M f(y).m_{r.x} d(y).$$

Then by the Arzela- Ascoli theorem there exists a sequence $(\ell_j)_j \subset N$ such that:

$$P_t f := \lim_{j\to\infty} \left(m_{\sqrt{2(n+2)t/\ell_j}} \right)^{\ell_j} f$$

Exists for all bounded $f \in C^{lip}(M)$, and it defines a Markov semi group on M satisfying $Lip(P_t f) \le e^{-kt} Lip(f)$ (see [140])

Theorem (1.1)[186]:

For any smooth complete Riemannian manifold M and any $K \in \mathbf{R}$, the following properties are equivalent:

(i) $Ric(M) \ge K$ which should be read as $Ric_x(\mathrm{V},\mathrm{V}) \ge K|\mathrm{V}|^2$ for all $x \in M, \mathrm{V} \in T_x M$.

(ii) The entropy $H(\cdot)$ is displacement K-convex on $\mathrm{P}^2(M)$.

(iii) the gradient flow $\Phi : \mathbb{R}_+ \times \mathrm{P}^2(M) \to \mathrm{P}^2(M)$ with respect to $H(\cdot)$ satisfies

$$\mathrm{W}_2(\Phi(\mathrm{t},\mu),\Phi(t,\mathrm{V})) \le e^{-kt}.\mathrm{W}_2(\mu,\mathrm{V}) \quad \forall \mu,\mathrm{V} \in \mathrm{P}^2(M)\ \forall t \ge 0.$$

Proof:

(ii) $\Rightarrow$ (i). Assume $\rightarrow$ (i). Then $Ric_0(e_1,e_1) \le K - \varepsilon$ for some $O \in M$, some unit vector $e_1 \in T_0 M$ and some $\varepsilon > 0$ let $e_1, e_2, \ldots, e_n$ be an orthonormal basis of $T_0 M$ such that:

$$R(e_1,e_i)e_1 = k_i e_i$$

For suitable number $k_i, i = 1,\ldots,n$ (denoting the sectional curvature of the plane spanned by e_1 and e_i if $i \ne 1$). Then $\sum_{i=1}^{n} K_i = Ric_0(e_1,e_i) \le K - \varepsilon$.

For $\delta, r > 0$ let $A_1 := B_\delta(\exp_0(re_1))$ and $A_0 := B_\delta(\exp_0(-re_1))$ be geodesic balls, and let:

$$A_{\frac{1}{2}} := \exp_0\left(\left\{y \in T_0M : \sum_{i=1}^{n}\left(\frac{y_i}{\partial i}\right)^2 \le 1\right\}\right)$$

With $\delta_i := \delta.\left(1 + r^2\left(K_i + \frac{\varepsilon}{2n}\right)/2\right)$. Choosing $\delta << r << 1$ we can achieve that $V_{\frac{1}{2}} \in A_{\frac{1}{2}}$ for each minimizing geodesic $V : [0,1] \to M$ with $V_0 \in A_0, V_1 \in A_1$.

Now let μ_0 and μ_1 be the normalized uniform distribution in $A_{\frac{1}{2}}$. then

$$H(\mu_0) = -\log volA_0 = -\log C_n - n\log\delta + O(\sigma^2)$$
$$H(\mu_1) = -\log volA_0 = -\log C_n - n\log\partial + O(\sigma^2).$$

With $C_n := vol(B_1)$ in $\mathbf{R}^n$, where as

$$H(V) = -\log volA_{\frac{1}{2}} = -\log C_n - \sum_{i=1}^{n}\log\delta_i + O(\delta^2)$$

$$= -\log C_n - n\log\delta - r^2\frac{\frac{\varepsilon}{2} + \sum_{i=1}^{n}K_i}{2} + O(r^4) + O(\delta^2)$$

$$\ge -\log C_n - n\log\delta - r^2\frac{K - \varepsilon/2}{2} + O(r^4) + O(\delta^2).$$

Since the optimal mass transport from μ_0 to μ_1 (with respect to W_2) is along geodesics of M the support of $\mu_{1/2}$ must be contained in the set $A_{1/2}$. Hence

$H(\mu_{1/2}) \ge H(V)$

and thus

$$H\left(\mu_{\frac{1}{2}}\right) - \frac{1}{2}H(\mu_0) - \frac{1}{2}H(\mu_1) \ge -\frac{K}{2}r^2 + \frac{\varepsilon}{4}r^4 + O(r^4) + O(\delta^2) > -\frac{K}{8}W_2(\mu_0,\mu_1)^2$$

for $\delta << r << 1$.

(i) $\Rightarrow$(ii). Here we closely follow the argumentation of [31,142,110,149,161,166] and use their notation. Assume that $Ric(M) \ge K$. We have to prove that

$$H(\mu_0) \le (1-t)H(\mu_0) + tH(\mu_1) - \frac{K}{2}t(1-t)\mathrm{W}_2(\mu_0,\mu_1)^2$$

For each geodesic $t \mapsto \mu_t$ in $(\mathrm{P}^2(M), W_2)$ and each $t \in [0,1]$. Without restriction, we may assume that μ_0 and μ_1 are absolutely continuous (otherwise the right-hand side is infinite). Hence there exists a unique geodesic connecting them. It is given as $\mu_t = (F_t)\mu_0$ where $F_t(x) = \exp(-t\,\nabla\Phi(x))$ with a suitable function Φ. Moreover, with

$J_t := \det d\, F_t(x)$ and $S(r) := \dfrac{\sin\left(\sqrt{\frac{K}{n-1}}.r\right)}{\sqrt{\frac{K}{n-1}}.r}$

Which should be read as $S(r) := \dfrac{\sinh\left(\sqrt{\frac{-K}{n-1}}.r\right)}{\sqrt{\frac{-K}{n-1}}.r}$

If $K < 0$ and as $S(r) = 1$ if $K = 0$ and with $\mathrm{V}_t(x,y)$ being the volume distortion coefficient of [31], we deduce

$$H(\mu_t) = H(\mu_0) - \int_M \log J_t(x)\,\mu_0(dx)$$

and thus

$$-H(\mu_t) + (1-t)H(\mu_0) + tH(\mu_1) = \int_M \log J_t(x)\,\mu_0(dx) - t\int_M \log J_t(x)\,\mu_0(dx)$$

$$\ge n\int \log\left[(1-t)\mathrm{V}_{1-t}(F_1(x),x)^{\frac{1}{n}} + t\mathrm{V}_t(x,F_1(x))^{\frac{1}{n}} J_1(x)^{\frac{1}{n}}\right]\mu_0 d(x) \;-t\int_M \log J_1(x)\,\mu_0(dx)$$

$$\ge n\int \log\left[(1-t)\left[\frac{S(1-t)d(F_1(x),x)}{S(d(F_1(x),x))}\right]^{1-\frac{1}{n}} + t\left[\frac{S(td(x,F_1(x)))}{S(d(F_1(x),x))}\right]^{1-\frac{1}{n}} J_1(x)^{\frac{1}{n}}\right]\mu_0(dx)$$

$$-t\int_M \log J_1(x)\,\mu_0(dx)$$

$$\ge (n-1)\int \left[(1-t)\log S((1-S)d(F_1(x),x)) + \log S(td(F_1(x),x)) - \log S(d(F_1(x),x))\right]\mu_0(dx)$$

$$\geq \frac{K}{2} t(1-t) \int d^2(F_1(x), x) \mu_0(dx) = \frac{K}{2} t(1-t) \mathrm{W}_2(\mu_0, \mu_1).$$

Here the first and second inequalities follow in [31,254,69,162,31,156,34,181,174]. The third inequality follows from the concavity of logarithm, and the least one from the fact that

$$(1-t)\log S((1-t)r) + t\log S(tr) - \log S(r) - \frac{t(1-t)}{2}\frac{K}{2}r^2$$
$$= (1-t)\lambda((1-t)r) + t\lambda(tr) - \lambda(r) \geq 0$$

for all $r \geq 0$ under consideration and $t \in [0,1]$ since $\lambda'(r) \leq 0$ where

$\lambda(r) := \log S(r) + \frac{1}{6}\frac{K}{n-1}r^2$. Note that according to the Bonnet-Myers theorem we may restrict ourselves to $r \geq 0$ with $\frac{K}{n-1}r^2 \leq \pi^2$.

In order to verify that $\lambda'(r) \leq 0$, it suffices to consider the cases $K = \pm(n-1)$. If $K = -(n-1)$, then $\lambda(r) = \log \sinh r - \log \frac{1}{6} r^2$ and $\lambda'(r) = \frac{\cosh r}{\sinh r} - \frac{1}{r} - \frac{1}{3} r$. The latter is non positive for all $r > 0$ if and only if $r\cosh r - \sinh r - \frac{1}{3} r^2 \sinh r \leq 0$ For all $r > 0$, Differentiating and dividing by $\frac{r}{3}$, we see that this is equivalent to $-r\cosh r + \sinh r \leq 0$, which (again by differentiation) will follow from $-r\sinh r \leq 0$ which is obviously true.

Analogously, if $K = n-1$ the condition $\lambda'(r) \leq 0$ is equivalent to $r\cos r - \sin r + \frac{1}{3} r^2 \sin r \leq 0$, which (by the same arguments as before) is equivalent to $-r \sin r \leq 0$. Here, of course, we have to restrict ourselves to $r \in [0, \pi]$.

Theorem (1.2) [186]: For any smooth complete Riemannian manifold M and $K \in \mathbf{R}$ the following properties are equivalent:

(i) $Ric(M) \geq K$.

(iv) For all $f \in C_c^\infty(M)$, all $x \in M$, and all $t > 0$, $|\nabla P_t f|(x) \leq e^{-kt} P_t |\nabla f|(x)$.

(v) For all $f \in C_c^\infty(M)$ and all $t > 0$, $\|\nabla P_t f\|_\infty \leq e^{-kt} \|\nabla f\|$.
For all bounded $f \in C^{lip}(M)$ and all $t > 0$, $Lip(P_t f) \leq e^{-kt} Lip(f)$.

Proof:

(i)$\Rightarrow$(iv) this is due to D.Bakry and M.Emery [28] and can be obtained using their Γ_2 calculus,(see [20]).

(iv) $\Rightarrow$ (v) take $\|.\|_\infty$ on both sides and use (on the right-hand side) the fact that P_t is a contraction on $L^\infty(M)$.

(v) $\Rightarrow$(i) we prove it by contradiction, assuming $\neg(i)\wedge(v)$. If (i) is not true, then there exists a point $O\in M$ and $\mathrm{V}\in \mathrm{S}^{n-1}\subset T_0M$ such that

$Ric_0(\mathrm{V},\mathrm{V})\le K-\varepsilon$ For some $\varepsilon>0$. If (v) is true then

$Lip(P_t f)\le e^{-kt}Lip(f)$

For all $f\in C_C^{Lip}(M)$ and all $t>0$. Indeed, fix, $y\in M$ and $t>0$, and choose $f_n\in C_C^\infty(M)$ with $f_n\to f$ uniformly on M and $Lip(f_n)\to Lip(f)$. Then

$|P_t f(x)-P_t f(y)|\leftarrow|P_t f_n(x)-P_t f_n(y)|\le e^{-kt}.d(x,y).Lip(f_n)\to e^{-kt}d(x,y).Lip(f)$.

Our first claim is that there exist a neighborhood U of O and a function $f\in C_C^{Lip}(M)$ such that $f|U\in C^2$ and

$$\nabla f(0)=0 \qquad (1)$$

$$Hess_0 f=0 \qquad (2)$$

$$|\nabla f(x)|=1 \qquad \forall x\in \mathrm{U} \qquad (3)$$

$$Lip(f)\le 1 \qquad (4)$$

In order to construct such a function f. Let $\mathrm{F}=\{x\in M\,|cut(0)|\log x\perp v\}\subset M$ be the orthogonal hyper surface to v in M and define the signed distance function f_0 from

F by

$f_0: M|cut(0)|\to\mathbf{R}$, $f_0(x)=dist(x,\mathrm{F}).\,sign(v,\log_0 x)$.

It is shown in Lemma (1.4) below that $f_0\in C^\infty(\mathrm{U}_0)$ for some neighborhood $\mathrm{U}_0\ni 0$ and that it satisfies the properties (1)-(4) from above with M (and U) replaced by U_0. Without restriction, we may assume that U_0 has a smooth boundary and compact closure. Now put $f=f_0\wedge P\vee(\approx P)$ with $P(x)=dist(x,M|\mathrm{U}_0)$. This function coincides with f_0 on a suitable neighborhood U of 0, and it has compact support and satisfies all the properties (1)-(4) from above.

Now let us fix a function f as above, choose atest function $0 \le \Phi \in C_C^\infty(\mathrm{U})$, and define

for $t>0$, $\Phi(t) := \int_m |\nabla p_t f|^2 \phi \, d\,vol$ .Then(v) and(4) imply

$$\int_m |\nabla ptf|^2 \phi dvol \le Lip(P_t f)^2 \int_M \phi dvol \le \exp(-2kt) \int_M \phi dvol = (1-2kt+O(t)) \int \phi dvol$$

Since $|\nabla P_t f|^2 \to |\nabla f|^2 = 1$ on $\operatorname{supp}(\Phi) \subset U$, the function Φextends continuously on the entire non negative half-line by $\phi(0) = \int_M \phi$ .by continuity of the function

$\Gamma_2(f,f) : U \to \mathbf{R}$.

$$\Gamma_2(f,f)(x) := \|Hess_x(f)\|^2 + Ric_x(\nabla f, \nabla f).$$

We find:

$$\Gamma_2(f,f) \le K - \frac{1}{2}\varepsilon$$

On some neighborhood of that contains, without loss of generality, U. From Bochner's formula we deduce.

$$\Phi(t) = \int_M 2(\nabla \rho_t \Delta f, \nabla \rho_t f) \phi dvol$$

$$= \int_M \left(2\Delta |\nabla \rho_t f|^2 - 2\Gamma_2(\rho_t f, \rho_t f)\right) \phi dvol = \int_M 2|\nabla \rho_t f|^2 \Delta\phi - 2\Gamma_2(\rho_t f, \rho_t f) \phi dvol \quad (5)$$

$$\xrightarrow{t \to 0} \int_M \left(2|\nabla f|^2 \Delta\phi - 2\Gamma_2(f,f)\phi\right) dvol$$

$$= -2\int_M \Gamma_2(f,f) \phi dvol \ge \left(\varepsilon - 2K\right) \int_M \phi = (\varepsilon - 2K)\Phi(0)$$

Thus $t \to \Phi(t)$ is differentiable in $t=0$, with $\Phi(0+) \ge (\varepsilon - 2K)\mathrm{E}(0)$. Consequently, we find for small t that

$$\Phi(t) \ge \Phi(0) + (\varepsilon - 2K)\Phi(0)t + O(t) = \Phi(0)(1 - 2Kt + O(t)), \text{ i.e.}$$

$$\int_M |\nabla P_t f|^2 \phi dvol \ge (1 + (\varepsilon - 2K)t + O(t)) \int_M \phi dvol$$

Which contradicts (5)

(vi) $\Rightarrow$ (v). This case is trivial.

One reason for the importance of Theorem (1.1) is that it characterizes lower Ricci bounds referring neither to the differential structure of M nor to the dimension of M. Property (ii) may be formulated in any metric measure space. For other weak substitutes of lower Ricci curvature bound including volume doubling and Poincare inequality, see [110, 142, 268, 216, 97, 49, 59, 27].

F.Otto and C. Villani [69] gave a very nice heuristic argument for the implication (i) $\Rightarrow$(ii). In the case $K = 0$, this implication was proven in [31,208,214,211,104.97].

The equivalence of (i) and (iv) is perhaps one of the most famous general results that relate heat kernels with Ricci curvature. It is due to [28,23,236,234,237,70,69,31], see also [20]. Property (iv) is successfully used in various applications as a replacement (or definition) of lower Ricci curvature bound for symmetric Markov semi groups on general state spaces. Our result states that (iv) can be weakened in two respects:

We can replace the point wise estimate by an estimate between L^∞ norm and one can drop the P_t on the right-hand side.

Besides being formally weaker than (iv), one other advantage of (v) is that it is an explicit (since P_t appears on both sides).

As an easy corollary to the equivalence of statements (iv) and (v), one may deduce the well-known fact that (iv) is equivalent to the assertion that for all f, x and t as above

$$|\nabla P_t f|(x) \leq e^{-kt}\left[P_t\left(|\nabla f|^2\right)(x)\right]^{\frac{1}{2}}$$

Property (vi) may be considered as a replacement (or as one possible definition) for lower Ricci curvature bounds for Markov semi group on metric spaces. For several non classical example (including nonlocal generators as well as infinite-dimensional or singular finite-dimensional state spaces) we refer to [76, 184, 140,144]. This property turned out to be the key ingredient to prove Lipschitz continuity for harmonic maps between metric spaces in [140,166,269,29,21].

According to the kantorovich-Rubinslein dually, property (vi) is equivalent to a contraction property for the heat kernels in terms of the L^1-Wasserstein distance W_1. Actually, however, much more can be proven.

Corollary (1.3) [186]: For any smooth complete Riemannian manifold M and any $K \in \mathbf{R}$, the following properties are equivalent:

(i) $Ric(M) \geq K$.

(vii) For all $x, y \in M$ and all $t > 0$, there exists $r \in [1, \infty]$ with:

$$W_r\left(P_t(x,\cdot), P_t(y,\cdot)\right) \leq e^{-kt}.d(x,y)$$

(viii) For all $r \in [1, \infty]$, all $\mu, v \in P^r(M)$, and all $t > 0$

$$W_r\left(\mu_{P_t}, V_{P_t}\right) \leq e^{-kt}.W_r(\mu, V).$$

(ix) For all $x_1, x_2 \in M$ there exists a probability space (Ω, A, P) and two conservative Brownian motions $(X_1(t))_{t\geq 0}$ and $(X_2(t)_{t\geq 0})$ defined on it with values in M and starting in x_1 and x_2 respectively, such that for all $t > 0$.

$$E\left[d\left(X_1(t), X_2(t)\right)\right] \leq e^{\frac{-kt}{2}}.d(x_1, x_2).$$

(x) There exists a conservative Markov process $(\Omega, A, P, X(t))_{x\in M\times M, t\geq 0}$ with values in $M \times M$ such that the coordinate processes $(X_1(t))_{t\geq 0}$ and $(X_2(t))_{t\geq 0}$ are Brownian motions on M and such that for all $x = (x_1, x_2) \in M \times M$ and all $t > 0$

$$d\left(X_1(t), X_2(t)\right) \leq e^{\frac{-kt}{2}}.d(x_1, x_2) P^x - a.s.$$

Proof: (vii) $\Rightarrow$ (vi). By Holder's inequality, property (vii) for $r \geq 1$ implies property (vii) for $r = 1$, which in true implies (vi) according to the Kantorovich Rubinstein dually.

Explicitly, for each coupling λ of $P_t(x,\cdot)$ and $P_t(y,\cdot)$

$$\left|p_t f(x) - p_t f(y)\right| = \left|\int\left[f(z) - f(w)\lambda(dzdw)\right]\right|$$

$$\leq Lip(f).\int d(z,w)\lambda(dzdw) \leq Lip(f).\left[\int d(z,w)^r \lambda(dzdw)\right]^{\frac{1}{r}}$$

Hence

$$\left|P_t f(x) - P_t f(y)\right| \leq Lip(f) d_r^w\left(P_t(x,\cdot), P_t(y,\cdot)\right) \leq Lip(f) d(x,y) e^{-kt}$$

(viii) $\Rightarrow$ (vii). Choose $\mu = \delta_x$ and $\upsilon = \delta_y$.

(ix) $\Rightarrow$ (vii). The distribution $\lambda(\cdot):=\mathrm{P}\left(X_1(2t),X_2(2t)\in\cdot\right)$ of the pair $(X_1(2t),X_2(2t))$ defines a coupling of $P_t(x_1,\cdot)$ and $P_t(x,\cdot)$. Hence

$$\mathrm{W}_1\left(P_t\left(x_1,\cdot\right).P_t\left(x_2,\cdot\right)\right)\leq\int d\left(z_1,z_2\right)\lambda\left(dz_1dz_2\right)=E\left[d\left(X_1\left(2t\right),X_2\left(2t\right)\right)\right]\leq e^{-kt}.d\left(x_1,x_2\right)$$

(x) $\Rightarrow$ (viii). Let λ be an optimal coupling of μ and V with respect to W_r, and let Π_t be the transition semi group of the Markov process from (x). Then

$\lambda_t:=\lambda\Pi_{2t}$ is a coupling of μ_{P_t} and v_{P_t}. Hence

$$\mathrm{W}_r\left(\mu_{p_t},\mathrm{V}_{p_t}\right)^r\leq\int d\left(w_1,w_2\right)^r\lambda_t\left(dw_1,dw_2\right)$$

$$=\int\int d\left(w_1,w_2\right)^r\Pi_{2t}\left(\left(x_1,x_2\right),dw_1dw_2\right)\lambda d\left(x_1d\,x_2\right)$$

$$=\int\mathrm{E}^{(x_1,x_2)}\left[d\left(X_1\left(2t\right),X_2\left(2t\right)\right)^r\right]\lambda\left(dx_1,x_2\right)$$

$$\leq e^{-ktr}.\int d\left(x_1,x_2\right)^r\lambda\left(dx_1,dx_2\right)=e^{-ktr}.\mathrm{W}_r\left(\mu,v\right)^r.$$

(x) $\Rightarrow$ (ix). Take expectations.

(i) $\Rightarrow$ (x). this implication is well-known and can be shown using either stochastic differential equation theory on Riemannian manifold in order to construct the coupling by a parallel transport process on $M\times M$ for two Brownian motions (cf.,[247,23,157] and [48,223,160]) or by a central limit theorem for coupled geodesic random walks and estimate of the type (2) (cf [184] for a similar argument).

Lemma (1.4)[186]:

Let M be a Riemannian manifold, $O\in M$, $\mathrm{V}\in T_0M$ and

$\mathrm{F}=\left\{\exp_0\left(u\right)\middle|u\in T_0M,u\perp\mathrm{V}\right\}\subset M$,the $(\mathrm{n}-1)$-dimensional hypersurface through

orthogonal to V. Then the signed distance function $f_0:M\to\mathbf{R}$ Belongs to c $^\infty(u)$ for some neighborhood $\mathrm{U}\ni o$ and $Hess_0(f_0)=0$.

Proof: the level sets $\mathrm{F}_\in=\left\{x\in M\,\middle|\,f_0\left(x\right)=\varepsilon\right\}$

Define a foliation of (a sufficiently small) neighborhood $\mathrm{u}\ni 0$ by smooth hyper surfaces. The unit normal vector field to F_ε is given by $\mathrm{V}=\nabla f_0$ which is well-defined and smooth sufficiently close to F (f_0 is a" distance function" on U in the sense of [213]. Hence the Hessian of f_0 in a point $p \in \mathrm{U}$ may be interpreted as the shape operator of F_ε in $p \in \mathrm{F}_\varepsilon$ with $\varepsilon = f_0(p)$,i.e.,

$$Hess_p(X,X) = \prod\nolimits_p^{\mathrm{F}_\varepsilon}(X,X) = \left\langle S_p^{\mathrm{F}_\varepsilon}(X), X \right\rangle_{T_pM}$$

Where $\prod_p^{\mathrm{F}_\varepsilon}$ is the second fundamental form of the hyper surface $\mathrm{F}_\varepsilon \subset M$ and $S^{\mathrm{F}_\varepsilon} : T_pM \to T_p\mathrm{F}_\varepsilon$ is the associated shape operator. The claim $Hess_0(f_0) = 0$ then follows from the construction of F, which implies that $\mathrm{F} = \mathrm{F}_0 \subset M$ is flat in 0, i.e., $S_0^{\mathrm{F}} = 0$

Theorem (1.5)[186]: For any smooth compact Riemannian manifold M and any $K \in \mathbf{R}$. The following properties are equivalent:

(i) $Ric(M) \geq K$.

(ii) The normalized Riemannian uniform distribution on spheres

$$\sigma_{r,x}(A) := \frac{\mathrm{H}^{n-1}(A \cap \partial B_r(x))}{\mathrm{H}^{n-1}(\partial Br(x))}, A \in \mathrm{B}(M),$$

Satisfies the asymptotic estimate

$$\mathrm{W}_2(\sigma_{r,x}, \sigma_{r,y}) \leq \left(1 - \frac{Kr^2}{2n} + O(r^2)\right).d(x,y) \qquad (6)$$

Where the error term is uniform with respect to $x, y \in M$.

(iii) The normalized Riemannian uniform distribution on balls

$$m_{r,x}(A) := \frac{m(A \cap B_r(x))}{m(B_r(x))}, \quad A \in \mathrm{B}(M)$$

Satisfies the asymptotic estimate

$$\mathrm{W}_r(m_{r,x}, m_{r,y}) \leq \left(1 - \frac{K}{2(n+2)} r^2 + O(r^2)\right).d(x,y) \qquad (7)$$

Where the error term is uniform with respect to $x, y \in M$.

Proof: $(xi) \Rightarrow (i)$. Let us define the family of Markov operators or: $\mathrm{F}_B \to \mathrm{F}_B$ by $\sigma_r f(x) = \int_M f(y)\sigma_{r,x}(dy)$ on the set f_b of bounded Borel- measurable functions on $M-$ using that for $f \in C^3(M)$

$$\sigma_r f(x) = f(x) + \frac{r^2}{2n}\Delta f(x) + O(r^2) \qquad (8)$$

Is given by $\mathrm{V} = \nabla f_0$ which is well-defined and smooth sufficiently close of F and on appropriate version of the Trotter –Chernov product formula [246,247,223,141,29]) applied to $p_t = \exp(t\Delta)$ as a Feller semi group on $\left(C(M), \|.\|_\infty\right)$, we find for all $f \in C(M)$

$$\left(\sigma_{\sqrt{2nt/j}}\right)^j f(x) \xrightarrow{j\to\infty} p_t f(x)$$

Uniformly in $x \in M$ and locally uniformly in $t \geq 0$ y the Rubinstein–Kantorovich duality condition, (xi) implies

$$\left|\sigma_r f(x) - \sigma_r f(y)\right| \leq \left(1 - \frac{k}{2n}r^2 + O(r^2)\right).d(x,y).Lip(f)$$

For all $f \in C^{Lip}(M)$ and $x, y \in M$, i.e.

$$Lip(\sigma_r f) \leq \left(1 - \frac{k}{2n}r^2 + O(r^2)\right).Lip(f)$$

and hence by interaction for $j \in N, r = \sqrt{2nt/j}$

$$Lip\left(\left(\sigma_{\sqrt{2nt/j}}\right)^j f\right) \leq \left(1 - \frac{kt}{j} + O\left(\frac{t}{j}\right)\right).Lip(f)$$

Passing to the limit for $j \to \infty$ yields

$$Lip(p_t f) \leq \exp(-kt).Lip(f)$$

Which is equivalent to (i) by Theorem (1.1).

For the proof of the converse we construct an explicit transport from $\sigma_{r,x}$ to $\sigma_{r,y}$ in the following lemma, whose proof is given below.

(i) $\Rightarrow$ (xi). We show this for the case $K < 0$; the case $K \geq 0$ is treated a negligible error $\Psi_r^{x,y}$ is under noting but parallel transport because

$$d\left(z,\Psi_r^{x,y}(z)\le d\left(z,\exp_y\left(\|_{V_{xy}}\log_x(z)\right)\right)+d\left(\exp_y\left(\|_V\log_x(z),\Psi_r^{x,y}(z)\right)\le d\left(z,\exp_y\left(\|_V\log_x(x)\right)\right)\right)\right)$$

$$+L\left\|\|_V\log_x(z)-\log_y\Psi^{x,y}(z)\right|\le d\left(z,\exp_y\left(\|_V\log_x(z)\right)\right)+Ld(x,y)O\left(r^2\right).$$

Where L is some uniform upper bound for the Lipschitz constant of $\log(\cdot)$ with respect to the second argument. The asymptotic inequality (6) is now easily verified from (8), since

$$W_1\left(\sigma_{r,x},\sigma_{r,y}\right)\le\frac{1}{H^{n-1}\left(\partial B_r(x)\right)}\int_{\partial B_r(x)}d\left(z,\Psi_r^{x,y}(z)\right)H^{n-1}d(z)$$

$$=\frac{1}{H^{n-1}\left(\partial B_r(x)\right)}\int_{\partial B_r(x)}d\left(z,\exp_y\|v\log_x z\right)H^{n-1}(dz)+d(x,y)O\left(r^2\right)$$

$$=d(x,y)+\frac{r^2}{2n}\Delta D^{x,y}(x)+d(x,y)O\left(r^2\right)$$

With $z\to D^{x,y}(z)=d\left(z,\exp_y\|v\log z\right)$, since

$$\Delta D^{x,y}(x)=trHess_x D^{x,y}=\sum_{i=2}^{n}I_M^v\left(J_i,J_i\right)$$

Where $I_M^v\left(J_i,J_i\right)$ is the index form of M along V_{xy} applied to the Jacobi field induced from parallel geodesic variations of V.

In the direction e_i with $\left\{\dot{V}_{xy}^{\varepsilon},e_2,\ldots,e_n\right\}$ being an orthonormal basic of T_xM. Hence we may conclude by the standard Ricci comparison argument that

$$\Delta D^{x,y}(x)\le 2(n-1)\sqrt{\frac{-K}{n-1}}\frac{\cosh\left(\sqrt{\frac{-K}{n-1}}d(x,y)\right)-1}{\sinh\left(\sqrt{\frac{-K}{n-1}}d(x,y)\right)}\le -Kd(x,y)$$

Such that we finally arrive at

$$W_1\left(\sigma_{r,x},\sigma_{r,y}\right)\le d(x,y)-\frac{r^2}{2n}Kd(x,y)+d(x,y)O\left(r^2\right).$$

(xii) $\Rightarrow$ (i). This is shown in the same way as the implication.

(xi) $\Rightarrow$ (i). With the slight difference that instead of (8) one uses

$$m_r f(x) = f(x) + \frac{r^2}{2(n+2)} \Delta f(x) + O(r^2).$$

(i) $\Rightarrow$ (xii). We proceed as before for (i) $\Rightarrow$ (xi) where, now we have to construct a map $\Phi_r^{x,y}$ that preserves the normalized uniform distributions on balls. However, since similarly to condition (16) in the proof of Lemma (1.6) we have

$$m_{r,x}(A) = \frac{1}{m(B_r(x))} \int_0^r \sigma_{u,x}(A) \mathrm{H}^{n-1}(\partial B_u(x))\,du$$

Such a map can be constructed from a map $\Psi_{r_1,r_2}^{x,y} = \partial B_r(x) \rightarrow \partial B_r(y)$ with

$\left(\Psi_{r_1,r_2}^{x,y}\right)_* \sigma_{r_1,x} = \sigma_{r_2,y}$ and that is almost induced from parallel transport in the sense of (9) below . It is clear that Lemma (1.6) can easily be generalized to yield such a map $\Psi_{r_1 r_2}^{x,y}$ which is all we need

Lemma (1.6)[186]:

Let M be a smooth compact Riemannian manifold and for $x \in M$ let $\sigma_{r,x}(\cdot)$ denotes the normalized Riemannian uniform distribution on $S_r(x) := \partial B_r(x)$ .then for r sufficiently small for each $x, y \in M$ there exists a geodesic segment $v = v_{xy}$ and a measurable map $\Psi_r^{x,y} : S_r(x) \rightarrow S_r(y)$ such that the push forward measure $\Psi_{r*}^{x,y} \sigma_{r.x}$ equals $\sigma_{r.y}$ and:

$$\sup_{z \in S_r(x)} \sup_{y \in M} \frac{\left|\log_x^z - \left\|_{\mathrm{V}}^{-1} \log_y \Psi_r^{x,y}(z)\right|}{d(x,y)} = o(r^2) \qquad (9)$$

Where the error term $o(r^2)$ is uniform in $x \in M$

Proof: We show the Lemma for the two-dimensional case first and inductively generalize this result to higher dimensions later. Let $n = 2$ and choose a parameterization of $S_r(x)$ and $S_r(y)$ (using Riemannian polar coordinates, for example) on $S' \subset R^2 \cong T_x M$, i.e., for all $f : S_r(x) \rightarrow \mathbf{R}$

$$\int_{S_r(x)} f(x) \mathrm{H}^{n-1} d(z) = \int_{S'} D_x\left(r, \mathrm{V}^0 \tilde{f}(\mathrm{V}^0) S(d\mathrm{V}^0)\right) = \int_0^{2\pi} D_x(r, S) \tilde{f}(S)\, ds$$

With a density $D_x(r, \mathrm{V}^0)$ given by:

$$D_x(t, \mathrm{V}^0) = \sqrt{\det\left(\left\langle Y_i(t, \mathrm{V}^0), Y_j(t, \mathrm{V}^0)\right\rangle\right)_{ij}}$$

$$t^{n-1}\left(1-\frac{t^2}{6}C_x\left(\mathrm{V}^0,\mathrm{V}^0\right)+O\left(t^2\right)\right) \quad (10)$$

Where C_x is the Gaussian curvature of (M,g) in x and $Y_j\left(t,\mathrm{V}^0\right)$ is the Jacobi field along $t\rightarrow \exp_x\left(t,\mathrm{V}^0\right)$ with $J_i(0)=0$ and $J_i'(0)=e_i$ for an orthonormal basis $\{e_i : i=1,2\}$ of T_xM;, for instance.

For x,y and $\gamma_{x,y}$ fixed, let the parameterization of $S_x'\subset T_xM$ and $S_y'\subset T_yM$ on $[0,2\pi]$ be chosen in such a way that $S_x'\ni 0\cong \dot{\gamma}_{xy}(0)$ and $S_y'\ni 0\cong \dot{\gamma}_{xy}(d(xy))$.

Next, we choose a function $\tau=\tau_r^{x,y}:[0,2\pi]\rightarrow[0,2\pi]$ with $\tau(0)=0$ satisfying

$$\frac{1}{\mathrm{H}^{n-1}\left(S_r(x)\right)}\int_0^u D_x(r,S)\,dx=\frac{1}{\mathrm{H}^{n-1}\left(S_r(y)\right)}\int_0^{\tau(u)} D_y(r,S)\,dS \quad (11)$$

For all $u\in[0,2\pi]$. Identifying $\tau:[0,2\pi]\rightarrow[0,2\pi]$ with the associated $\tau:S'\rightarrow S'$, then equation (11) just means that the induced map $\Psi=\Psi^{x,y}:S_r(x)\rightarrow S_r(y)$

$$\Psi(z)=\exp_x\left(r\tau\left(\frac{1}{r}\log_x(z)\right)\right)$$

Transport the measure $H_{r,x}^{n-1}$ into $H_{r,y}^{n-1}$. By the definition of $\Psi^{x,y}$, estimate (9) is equivalent to

$$\sup_{z\in[0,2\pi]}\ \sup_{y\in M}\frac{\left|\tau_r^{x,y}(z)-z\right|}{d(x,y)}=O(r) \quad (12)$$

Denoting:

$$E_x(r,S):=\frac{D_x(r,S)}{\mathrm{H}^{n-1}\left(S_r(x)\right)}=1+O\left(r^2\right) \quad (13)$$

With $E(\cdot,S)\in C^2(M\times[0,\varepsilon])$ for all $S\in[0,2\pi]$ and some $\varepsilon>0$ (11) yields

$$\left(1+O\left(r^2\right)\right)\left|\tau(z)-z\right|=\left|\int_\tau^{\tau(z)} E_y(r,S)\,ds\right|=\left|\int_0^z\left(E_x(r,S)\right)-E_y(r,S)\,dS\right|.$$

Consequently.

$$\sup_{y\in M}\frac{\left|\tau_r^{x,y}(z)-z\right|}{d(x,y)}\le\left(1+O\left(r^2\right)\right)\int_0^z\left|\nabla_x E(r,S)\right|ds$$

Due to $E(0,S)\equiv 1=const$, we find $\lim_{r\to 0}|\nabla_x E(r,S)|=0$ and hence

$$\lim_{r\to 0}\sup_{y\in M}\frac{|\tau_r^{x,y}(z)-z|}{rd(x,y)}\le \lim_{r\to 0}\int\frac{\partial}{\partial r}|\nabla_x E(r,S)|dS$$

$$=\lim_{r\to 0}\int_0^z\left\langle\frac{\nabla_x E(r,S)}{|\nabla_x E(r,S)|},\frac{\partial}{\partial r}\nabla_x E(r,S)\right\rangle_{T_xM} dS=\lim_{r\to 0}\int_0^z\left\langle\frac{\nabla_x E(r,S)}{|\nabla_x E(r,S)|},\nabla_x\frac{\partial}{\partial r}E(r,S)\right\rangle_{T_xM} dS \quad (14)$$

The right –hand side of (13) yields

$$\frac{\partial}{\partial r}E(r,S)\big|_{r=0}=0$$

From which we see that the integral above vanishes for r tending to 0. this establishes (12) for fixed x. By the smoothness of (M,g) the error term is locally uniform in $x\in M$ and hence is also globally uniform since M is compact. See [145, 149, 144, 142, 193, 133, 31]

The case $\dim M=3$ will show how we can deal will arbitrary dimensions $n\in N$. Fix $x,y\in M$ as well as some segment V_{xy} from x to y. By means of the inverse of the exponential map, we lift the measures $\sigma_{r,x}$ and $\sigma_{r,y}$ onto the unit sphere in T_xM and T_yM, respectively, which we disintegrate along the $\dot v_{xy}$-direction as follows: choose an orthonormal basis $\{e_1,e_2,e_3\}\subset T_xM$ with $e_1=\dot V_{xy}$ and $\{e_1',e_2',e_3'\}=\|_V\{e_1,e_2,e_3\}\subset T_yM$, and for $u\in[-1,1]$ let $S_{r,x}^{\perp}(u)=\{\exp_x(r(ue_1,se_2,te_3))|u^2+s^2+t^2=1\}$ denote the “orthogonal” part of $S_r(x)$ with respect to V at ue_1. Define a probability measure $C_{r,x}(u)(du)$ on $[-1,1]\subset e_1R\subset T_xM$ by

$$C_{r,x}(u)(du)=\frac{1}{H^{n-1}(S_r(x))}H^{n-2}(S_{r,x}^{\perp}(u)du$$

$$=\frac{1}{\int_{S^2}D_x(r,v^0)S^2(dv^0)}\int_{S^1_{\sqrt{1-u^2}}}D_x(r,(u,\eta))S^1(d\eta)du$$

With the Riemannian volume density

$$D_x(t,V^0)=t^{n-1}\left(1-\frac{t^2}{6}Ric_x(V^0,V^0)+(t^2)\right) \quad (15)$$

and $C_{r,y}(u)$ analogously, let $\tau_1=[-1,1]e_1\to[-1,1]e_1'$ be the function defined by

$$\int C_{r,x}(u)\,du = \int_0^{\tau_1(x)} C_{r,y}(u)\,du \qquad \forall S \in [-1,1] \tag{16}$$

For each $S \in [-1,1]$ define a transport

$$\tau_S^{\perp} : S_{r,x}^{\perp}(S) \to S_{r,y}^{\perp}(\tau_1(S))$$

Analogously to equation (11) preserving the probability measure $\sigma_{r,x}^{\perp}(S)(\cdot)$ and $\sigma_{r,y}^{\perp}(\tau(S))(\cdot)$ obtained from conditioning $\sigma_{r,x}$ and $\sigma_{r,y}$ on $S_{r,x}^{\perp}(S)$ and $S_{r,y}^{\perp}(\tau_1(S))$, respectively. Hence the map

$$\Psi^{x,y} : S_r(x) \to S_r(y), \Psi^{x,y}(z) := \exp_y\left(rt\left(\frac{1}{r}\log_x(z) \right) \right)$$

induced from $\tau = \tau_r^{x,y} : T_x M \supset_{\perp}^{2} (x) \to S^2(y) \subset T_y M$

$\tau(u) = \left(\tau_1(u_1), \tau_{u_1}^{\perp}((u_2, u_3))\right)$.

Will push forward $\sigma_r(x)$ into $\sigma_r(y)$, and it remains to prove the asymptotic estimate (12). Since the distance $\left|\tau_r^{x,y}(u) - u\right|^2$ is Euclidean, we may use estimate (12) from the two-dimensional case for the $\left|\tau_r^{x,y}\left((u_2, u_3)\right) - (u_2, u_3)\right|$ part, which also persists in this situation. Indeed, it is sufficient to note that expression (14) and hence the error estimate

$$\sup_{z \in [0,2\pi]} \sup_{y \in M} \frac{\left|\tau_{x,y}^{\perp}((u_2, u_3)) - (u_2, u_3)\right|}{d(x,y)} = O(r)$$

Also hold true for the embedded orthogonal spheres $S_{r,x}^{\perp}(u)$ and $S_{r,y}^{\perp}$ (since they are parallel translates of one another) and to note that by the triangle inequality, this also generalizes to the situation $\tau_{x,y}^{\perp} : S_{r_1}^{\perp}(x) \to S_{r_2}^{\perp}(y)$ with $r_1, r_2 \le r$. Thus it remains to prove

$$\sup_{u_1 \in [-1,1]} \sup_{y \in M} \frac{\left|\tau_1 u_1 - u_1\right|}{d(x,y)} = O(r)$$

which follows from (16) by argument similar to those that established (12) in the two-dimensional case. This completes the proof in three dimensions. For arbitrary $n \in N$, one proceeds in a similar fashion by inductively reducing the problem to lower dimensions.

Chapter 2

Free Probability of Wasserstein metric On Trace-State Space

In dimension one, we prove that the square of the Wasserstein distance to the semi-circle distribution is majorized by a manifold free entropy quantity.

The Wasserstein distance between two probability distributions μ, V on $\mathbf{R}^n$ is given by

$$\mathrm{W}(\mu,\mathrm{V}) = \inf_{\pi\in\Pi(\mu,\mathrm{V})}\left(\int |x-y|^2\, d\pi(x,y)\right)^{1/2} \tag{1}$$

Where denotes the probability measures on. $\mathbf{R}^n \times \mathbf{R}^n$ with marginals μ and V .Following the usual free probability recipe we shall replace the set of probability measures by the trace-state space of a C^* -algebra and take margillals with respect to a free product.

We note that in the context of non-commutative geometry, there is a different non-commutative extension, due to A. Connes [3], of the related Monge-Kantorowitz metric. The Monge-Kantorowitz metric is a $p=1,\ p-$ Wasserstein metric, but the definition which is extended is the dual definition based on Lipschitz functions, and the extension involves Fredholmmodules or derivations (see [183, 200, 198, 44, 27,39]We will work in the framework of tracial C^*-probability spaces (M,τ), where M is a unital C^* -algebra and τ is a trace state. The simplest is to define the metric at the' level of noncommutative. Random variable . If $(X_1,...,X_n)$ and $(Y_1,...,\mathrm{Y}_n)$ are two tuples of noncommutative random variables in tracial C^* -probability spaces (M_I,τ_I) and (M_2,τ_2), we define as the infimum of

$$\left\|\left(\left|X'_j - Y'_j\right|_p\right)_{1\le j\le n}\right\|_p \tag{2}$$

over $2n$ -tuples $(X'_1,...,X'_n,Y'_1,...,Y'_n)$ of noncommutative random variables in some tracial C^*-probability space (M_3,τ_3) such that the n -tuples $(X'_1,...,X'_n),(X_I,...,X_n)$ and respectively $(Y'_1,...,Y'_n)(X_1,...,X_n,)$ have the same $*$-distributions. Here $|\cdot|_p$ is the p-norm in a tracial C*-probability space. while $\|\cdot\|_p$. lip is the p -norrn on $\mathbf{R}^n$. Like in the classical case, if $p=2$ we call W_p the free Wasserstein metric' and we will. also use the notation W for W_2. We shall refer to W_p as the free p- Wasserstein metric. Note also that if

$$X_j = D_j + iE_j,\ Y_j = F_j + iG_j, \tag{3}$$

where D_j, E_j, F_j, G_j are self-adjoint, then

$$\mathrm{W}\left((X_I,\ldots,X_n),(\mathrm{Y}_1,\ldots,Y_n)\right) = \mathrm{W}\left((D_I,\ldots,D_n,E_I,\ldots,E_n),\ (F_I\ldots,F_n,G_I,\ldots,G_n)\right) \tag{4}$$

Note also that $W_p\left((X_I,\ldots,X_n),(Y_I,\ldots,Y_n)\right)$ depends only on the $*$-distributions of $.(X_1\ldots,X_n)$and$(Y_I,\ldots,Y_n)$ If we consider n-tuples with the same *-distribution as equivalent; then W_p will be a distance between equivalence' classes of $n-$ tuples. We pass now to trace-state spaces $TS(A)$, where A is a unital C^{*-}algebra. We will assume A is finitely generated and we will assume such a generator $(a_l,\ldots,a_n)$ has been specified. The p-Wasserstein metric on $TS(A)$ given by

$$\mathrm{W}_p(\tau',\tau'') = \mathrm{W}_p\left((a_1',\ldots,a_n'),(a_1'',\ldots,a_n'')\right), \tag{5}$$

where $\tau',\tau'' \in TS(A)$ and $(a_1',\ldots,a_n')$ and $(a_1'',\ldots,a_n'')$ demote the variables defined by $(a_1.\,..,a_n)\mathrm{in}(A,\tau')$ and respectively (A,τ'').

This definition can be rephrased using free products. If A_I, A_2 are until C^*-algebras, we denote by $\sigma_j : A_j \to A_l * A_2$ the canonical injection of A_j into the full free product C^*-algebra (this presumes amalgamation over 0). If $\tau_j \in TS(A),(1 \le j \le 2)$we define

$$TS(A_l{*}A_2;\tau_1,\tau_2) = \left\{\tau \in TS\left(A_l * A_2\right)\tau\circ\sigma_j = \tau_j,\ j=1,2\right\}. \tag{6}$$

Remark that $\tau_1 * \tau_2 \in TS(A_l * A_2;\tau_1,\tau_2)$ .It is easy to see that

$$\mathrm{W}_p\tau',\tau'' = \inf\left\{\left\|\left(\left|\sigma_1\left(a_j\right)-\sigma_2\left(a_j\right)\right|_{p,\tau}\right)_{1\le j\le n}\right\|_p \left|\tau \in TS\,(A * A;\tau',\tau'')\right.\right\}, \tag{7}$$

Where denotes the p-norm in $L^p(A;\tau)$.Remark also that the distance on n-tuples of variables can be obtained from the definition for trace-states. Assume for simplicity $X_j = X_j^*,\ Y_j = Y_j^*$ and $\mathbf{R} \ge \|X_j\|, \mathbf{R} \ge \|Y_j\|,\ 1 \le j \le n$. Let then $A = \left(C\left|-\mathbf{R},\mathbf{R}\right|\right)^{*n}$ (the free product of n copies) and $\sigma_k(a) = a_k$, where a is the identical function in $C[-\mathbf{R},\,\mathbf{R}]$. Let $\rho_j : A \to M_j,\ j=1,2$ be the *-homomorphisms such that $\rho_j(a_k) = X_k,\ \rho_2(a_k) = Y_k$ where the X_k's are in (M_l,τ_I) and the Y_k's in (M_2,τ_2). Then

$$\mathrm{W}_p(\tau',\tau'') = \mathrm{W}_p\left((X_I,\ldots,X_n),(Y_1,\ldots,Y_n)\right), \tag{8}$$

where $\tau' = \tau_1 \circ \rho_1, \tau'' = \tau_2 \circ \rho_2$.

Theorem (2.1)[198]: W_p is a metric:.

Proof: To check that W_p is a metric on the set of equivalence classes of n-tuples of variables or equivalently on a trace-state space $ST(A)$ from equation (5),the nontrivial assertion is the triangle inequality. Indeed that

$W_p((X_1,...,X_n),(Y_I,...,Y_n))=0 \Leftrightarrow (X_I,...,X_n),(Y_1,...,Y_n)$ have the same *-distribution or

$W_p(\tau',\tau'')=0 \Leftrightarrow \tau'=\tau''$ are easy to see. For the triangle inequality it will suffice to prove it in the equation (2).

Let $(X_1',...,X_n',Y_1',...,Y_n')$ in (M_{12},τ_{12}) and $(Y_1'',...,Y_n'', Z_1'',...,Z_n'')$ in $(M_{23}.\tau_{23})$ be $2n$-tnples in tracial W^*-probability spaces such that $(X_1',...,X_n'):(X_I,...,X_n)$, $(Y_1',...,Y_n'):(Y_1'',...,Y_n'')\sim(Y_I,...,Y_n)$, $(Z_1'',...,Z_n'')(Z_1,...,Z_n)$

Where : the –tuples have equal *-distribution. There is a trace- preserving automorphism of $W^*(Y_1',...,Y_n')$ and $W^*(Y_1'',...,Y_n'')$ which identifies Y_j' and Y_j''. Abusing notation we shall denote by M_2 the von Neumann sub algebras of M_{I2} and M_{23} generated by $(Y_1',...,Y_n')$ and respectively $(Y_1'',...,Y_n'')$ identified as above. Let E' and E'' be the conditional expectations of M_{12} and respectively M_{23} onto M_2. Let $(M_{123},E)=(M_{12}E')*_{M_2}(M_{23}E'')$ and $\tau_{123}=\tau_2\circ E$ where $\tau_2=\tau_{12}|M_2=\tau_{23}|M_2$ (see in [42,52,196]). Further, with $\rho_{12}:M_{12}\to M_{123}$, $\rho_{23}:M_{23}\to M_{123}$ denoting the canonical embeddings, let $X_j'''=\rho_{12}(X_j'), Z_j'''=Z_j''$ Then $\rho_{12}(Y_j')=\rho_{23}(Y_j'')$ implies.

$$\left|X_j'''-Z_j'''\right|_{p,\tau_{123}} \le \left|X_j'''-\rho_{12}(Y_j')\right|_{p,\tau_{123}}+\left|\rho_{23}(Y_j'')-Z_j'''\right|_{p,\tau_{123}}=\left|X_j'-Y_j'\right|_{p,\tau_{12}}+\left|Y_j''-Z_j''\right|_{p,\tau_{23}}$$

which is precisely what we need to establish the triangle inequality

$$W_p((X_I,...,X_n),(Y_1,...,Y_n))+W_p((Y_1,...,Y_n),(Z_1,...,Z_n))\ge W_p((X_I,...,X_n),(Z_1,...,Z_n))$$

Let as also record as a proposition some easy consequences of the capacity of the trace-state space.

Proposition (2.2)[198]: (a) The infimum in the definition of W_p is attained (both in the equation (2) and (8)).

(b) Let $\tau_1^k,\tau_1,\tau_2^k,\tau_2 \in TS(A)$ and assume $\tau_2^k)$ converges weakly to τ_j as

$k \to \infty \left(j = 1.2 \right)$

Then

$$\liminf_{k\to\infty} \ \mathrm{W_p}\left(\ \tau_1^{(k)}, \tau_2^{(k)} \right) \geq \mathrm{W_p}\left(\tau_1, \tau_2\right). \qquad (9)$$

Let $\left(X_1^{(k)},...,X_n^{(k)}\right), \left(X_I,...,X_n\right), \left(Y_1^{(k)},...,Y_n^{(k)}\right)\left(\mathrm{Y}_1,...,Y_n\right))$ be n-tuples of' variables in tracial C^* -probability spaces and assume that $\left\|X_j^{(k)}\right\| \leq \mathbf{R}, \left\|X_j\right\| \leq \mathbf{R}, \left\|Y_j^{(k)}\right\| \leq \mathbf{R}, \left\|Y_j\right\| \leq \mathbf{R},$ and

that $\left(X_1^{(k)},...,X_1^{(k)}\right)$, $\left(Y_1^{(k)},...,Y_1^{(k)}\right)$ converge in *-distribution to $\left(X_1,...,X_n\right)$ and respectively $\left(Y_1,...,Y_n\right)$. Then

$$\liminf_{k\to\infty} \mathrm{W}_p\left(\left(X_1^{(k)},...,X_n^{(k)}\right),\left(Y_1^{(k)},...,Y_n^{(k)}\right)\right) \geq \mathrm{W}_p\left(\left(X_I,...,X_n\right),\left(\mathrm{Y}_1,...,Y_n\right)\right). \qquad (10)$$

If $\left(X_I,...,X_n\right)$ are commuting self-adjoins variables in a tracial C^*-- probability space, then their distribution $\mu_{X_1,...,X_n}$ is a compactly supported probability measure on $\mathbf{R}^n$.

Theorem (2.3)[198]: Let $\left(X_I,...,X_n\right)$ and $\left(\mathrm{Y}_1,...,Y_n\right)$ be two n-tuples of commuting self-adjoint variables in tracial C^*-probability spaces. Then the free and classical Wasserstein distances are equal: $\mathrm{W}\left(X_I,...,X_n\right)\left(\mathrm{Y}_1,...,Y_n\right) = \mathrm{W}\left(\mu_{x_1,...,x_n}, \mu_{Y_1,...,Y_n}\right)$.

Proof: The left-hand side is the right-hand side, since the classical Wasserstein distance can be defined the same way as the free one, with the only difference that the 2n-tuples $\mathrm{X}_1',...,\mathrm{X}_n', \mathrm{Y}_1',...,\mathrm{Y}_n')$ in the infimum are required to live in commutative tracial C^*-probability spaces.

Let $(X_1', ..., X_n', \mathrm{Y}_1', ..., Y_n')$ be a 2n-tuple in the infimum defining the free distance. Passing to the van Neumann algebra completion, we may assume (M_3, τ_3), where X_j', Y_j' live, is a W^*-probability space with a normal faithful trace state. Let $A = \mathrm{W}^*(X_1',...,X_n') \subset M_3$, $B = \mathrm{W}^*(Y_1',...,Y_n') \subset M_3$ and let E_A be the canonical conditional expectation onto A. Then the unital trace-preserving completely positive map $\phi = \mathrm{E}_A \left| B : B \to A \right.$ gives rise to a state $\mathrm{V} : A \otimes B \to \mathrm{C}$**,** on a commutative algebra, defined by

$$\mathrm{V}\left(a \otimes b\right) = \tau_3\left(a\phi\left(b\right)\right).$$

The positively of V.

$$\tau_3\left(\sum_{i,j} a_i a_j^* \phi\left(b_i b_j^*\right)\right) \geq 0$$

is easily inferred from the positivity of the matrix $\left(\phi\left(b_i b_j^*\right)\right)_{i,j}$ Alternatively, probabilistically, V. is the probability measure on $\mathbf{R}^{2n}$ obtained by integrating $\phi : L^\infty\left(\mathbf{R}^n, \mu_{Y_1,\ldots,Y_n}\right) \to L^\infty\left(\mathbf{R}^n, \mu_{X_1,\ldots,X_n}\right)$.Then

$$\sum_{1\leq j\leq n} \mathrm{V}\left(\left(X_j' \otimes I - I \otimes Y_j'\right)^2\right) = \sum_{1\leq j\leq n} \tau_3\left(X_j'^2 + \phi\left(Y_j'^2\right) - 2X_j'.\phi\left(Y_j'\right)\right)$$

$$= \sum_{1\leq j\leq n} \tau_3\left(X_j' + \left(Y_j'^2\right) - 2\mathrm{E}_{Aj}\left(X_j' Y_j'\right)\right) = \sum_{1\leq j\leq n} \tau_3\left(\left(X_j' - Y_j'\right)^2\right).$$

Sine $A \otimes B$ is commutative this proves the theorem.

Let X, S in (M, τ) be self-adjoint and freely independent and' assume S is $(0,1)$ semi circular. The purpose is to estimate $\mathrm{W}(X,S)$. We begin by studying variables $X(\mathrm{t}) = \mathrm{e}^{-\mathrm{t}/2}X + \left(1-\mathrm{e}^{-\mathrm{t}}\right)^{1/2} S$ which have the same distribution as the variables in the free Ornstein-Uhlenbeck process. For technical reasons, and without extra work, the complex PDE will be derived under the more general assumption that X is unbounded self-adjoint affiliated with M (see [92]).If Y is self-adjoint affiliated with M, we denote by μ_X its distribution and by $G\mu_Y(z)$ or $G_Y(z)$ the Cauchy transform of μ_Y, which equals $\tau\left(\left(zI - X\right)^{-1}\right)$.Let $\tilde{G}(r,z) = G_{Y(r)}(z)$ and $\mathrm{G}(\mathrm{t}, \mathrm{z}) = G_{X(t)}(\mathrm{z})$, If $Y(r) = X + r^{1/2}S$, $\mathrm{Im} \approx> 0, r > 0, t \geq 0$.Then $\tilde{G}$ satisfies the complex: .Burgers equation (see[201],[47])

$$\frac{\partial \tilde{G}}{\partial r} + \tilde{G}\frac{\partial \tilde{G}}{\partial z} = 0$$

Like $\tilde{G}(t,z)$ also $G(t,z)$ is C^1 on $[0, \infty) \times \left\{\mathrm{z} \in \mathrm{C} \,\middle|\, \mathrm{Im}\, z > 0\right\}$ and holomorphic in z for fixed t.

Note that $\mathrm{X}(\mathrm{t}) = \mathrm{e}^{-\mathrm{t}/2}\mathrm{Y}\left(\mathrm{e}^{\mathrm{t}}\right)$ and that $\mathrm{G}_\alpha Y(\mathrm{z}) = \alpha^{-1}\mathrm{G}\left(\alpha^{-1}z\right)$.It follows that $\mathrm{G}(\mathrm{t,z}) = \mathrm{e}^{t/2}\tilde{\mathrm{G}}\left(e^t, e^{t/2}z\right)$, complex Burgers equation then gives

$$\frac{\partial G}{\partial t} + \left(G - \frac{z}{2}\right)\frac{\partial G}{\partial z} - \frac{1}{2}G = 0 \qquad (11)$$

With initial data $\mathrm{G}(0,\mathrm{z}) = \mathrm{G}_\mathrm{x}(\mathrm{z})$.

Here we shall assume that the distribution of X is of the form $p_\lambda * \mu$ where p_λ is the Cauchy distribution with density $\pi^{-1}\lambda\left(\lambda^2+x^2\right)^{-1}(\lambda>0)$ and μ has compact support. Since $p_\lambda * \mu = p_\lambda (\ \mu \ (see[92])$ this is equivalent to replacing X with $X+\lambda C$ where X is bounded, X and C are free and C has a Cauchy distribution p_1.Note that,

$$\mu_X+\lambda C+r^{1/2}S=\mu_X+r^{1/2}S*p_\lambda, G_{X+\lambda C+r^{1/2}S}(z)=G_{X+r^{1/2}S}(z+i\lambda)$$

etc. Thus, if the distribution of X is of the form $p_\lambda * \mu$ then the equation (11) is satisfied on an extended domain $\left\{(t,z)\in[0,\infty)\times\mathrm{C}\,\middle|\,\mathrm{Im}\,z>-e^{t/2}\lambda\right\}$

Let $-\pi^{-1}G(x,t)=q(x,t)+ip(x,t)$ where $x\in\mathbf{R}$ then $p(.,t)$ is the density of $\mu_X(t)$ and is analytic. For fixed t and $k\geq 0$ We have

$$\left|\frac{\partial^k}{\partial x^k}p(x,t)\right|=O\left(\left(1+|x|\right)^{-2-k}\right) \text{ and } \left|\frac{\partial^k}{\partial x^k}q(x,t)\right|=O\left(\left(1+|x|\right)^{-1-k}\right).$$

Moreover these bounds are uniform for t in a compact set. Equation (12) gives

$$\left.\begin{aligned} q_t &= \pi(qq_x - \mathrm{pp}_x)+\ 2^{-1}(xq_x+q) \\ P_t &= \pi(pq_x+qp_x)+2^{-1}(xp_x+p) \\ q &= -Hp, \end{aligned}\right\} \qquad (12)$$

Where H denotes the Hilbert transform.

Since $p(x,t)>0$ we infer that $f(a,t)=\int_{-\infty}^{\alpha}p(x,t)dx$ is a C^∞-diffeomorphisms $f(.t):\mathbf{R}\to(0,1)$ which transports $\mu_X(t)$ to Lebesgne measure. Hence $\phi_{s,t}(.)=f^{-1}(f(.\,s),t)(0<\mathrm{s}<t)$ will be a C^∞-diffeomorphisms $\mathbf{R}\to\mathbf{R}$ which transports $\mu_X(s)$ to $\mu_X(t)$. This is the same as saying that $X(t)$ and $\phi_{s,t}(X(s))$ have the same distribution. It is easily seen that

$$\frac{\partial}{\partial t}f^{-1}(y,t)=\frac{-\left(\frac{\partial}{\partial t}f\right)\left(f^{-1}(y,t),t\right)}{p\left(f^{-1}(y,t),t\right)}$$

Using (12) to compute $\frac{\partial}{\partial t}f$ we find

$$\frac{\partial}{\partial t}f(a,t)=\int_{-\infty}^{\alpha}\left(\pi(pq)_x+2^{-1}(xp)_x\right)dx=\pi(pq)(a,t)+2^{-1}ap(a,t).$$

$$\frac{\partial}{\partial t} f^{-1}(y,t) = \pi q\left(f^{-1}(y,t),t\right) - 2^{-1} f^{-1}(y,t).$$

For $y = f(\mathrm{x,s})$ we get the transport equation .

$$\frac{\partial}{\partial t}\phi_{s,t}(x) = \pi Hp(.,t)\left(\phi_{s,t}(x)\right) - 2^{-1}\phi_{s,t}(x) \tag{13}$$

with initial condition $\phi_{s,s}(x) = x$.

By the L^m-continuity $(1 < m < x)$ results for the density (see [47]) applied to $\mu($ $\mu_{r^{1/2}S}$ as a function of r, we infer after convolutions with Cauchy distributions the continuity of $(0,\infty) \circ\ t \to Hp(.,t) \in L^m(\mathbf{R})$

(the L^m-space w.r.t. Lebesgue measure). We should keep these facts in mind in computations where we shall use (13).

Lemma (2.4)[198]: Assume X has distribution $\mu * p_\lambda$ where μ. has compact

support and let $X(t) = e^{-t/2} X +\ (1-e^{-t})^{I/2} S$ with $\mathrm{S}\ (0,\ \mathrm{I})-$semicircular and free from X. Let $g \in \mathrm{C}^\infty(\mathbf{R})$ be such that $\|g\|_\infty < \infty, \|g'\|_\infty \le 1$ and assume g' has compact support. Then

$$(\mathrm{t}\ -\ \mathrm{s})^2\ \mathrm{W}\left(g\left(\mathrm{X}(s)\right), g\left(\mathrm{X}(\mathrm{t})\right)\right)^2 \le \sup_{s \le h \le t} \int_{\mathrm{supp\,g}'} \left(\pi Hp(.,\ \mathrm{h})(x\ ) - 2^{-1}\mathrm{x}\right)^2 \mathrm{p}(\mathrm{x,h})\mathrm{dx}.$$

Proof: We have

$$\mathrm{W}\left(g\left(X(\mathrm{s})\right), g\left(X(\mathrm{t})\right)\right)^2 \le \int_{\mathbf{R}} \left|g(x) -\ g\phi\left(\phi_{s,t}(x)\right)\right|^2 p(x,\mathrm{s})\,dx$$

$$\le \int_{\mathbf{R}} \Big(\int_S^1 g'\left(\phi_{s,h}(x)\right)(\pi Hp(.,h)\left(\phi_{s,h}(x)\right) - 2^{-1}\phi_{s,h}(x))\,dh\Big)^2 p(x,s)\,dx$$

$$\le (t-s)\int_S^1 \Big(\int_{\mathbf{R}} \left(g'\left(\phi_{s,h}(x)\right)\right)^2 (\pi Hp(.,h)\left(\phi_{s,h}(x)\right) - 2^{-1}\phi_{s,h}(x))^2\, dhp(x,s)\,dx\Big)dh$$

$$= (t-s)\int_S^t \left(\int_{\mathbf{R}} \left(g'\left(\phi_{s,h}(x)\right)\right)^2 \left(\pi H_p(.,h)\left(\phi_{s,h}(x)\right) - 2^{-1}\phi_{s,h}(x)\right)^2 dhp(x,s)\,dx\right)$$

$$= (t-s)\int_S^1 \int_{\mathbf{R}} \left(g'(x)\right)^2 (\pi Hp(.,h)(x) - 2^{-1}x)^2\, p(x,h)\,dxdh$$

$$\le (t\ -s\ )^2 \sup_{s \le h \le t} \int_{\mathrm{supp\,g}'} (\pi Hp(.,h)(x\ ) - 2^{-1}x\ )^2\, p(x\ ,h)\,dx$$

Assume X is bounded and the semicircular variable S is free W.r.t. X. Then the distribution $\mu_X(t)$ of $X(t)=e^{-1/2}X+\left(1-e^{t}\right)^{1/2}S$ has L^{∞}-density $p(.,\mathrm{t})$ w.r.t Lebesgue measure (see any of the papers [92], [203,201], [45,47]).

Lemma (2.5)[198]: Assume X is bounded, S is $(0,1)$ semicircular, X and S are free and let $p(.,t)$ be the density of $\mu_X(\mathrm{t})$, where $X(t)=e^{-1/2}X+\left(1-e^{-t}\right)^{1/2}S$. Then

$$\leq (t-s)^{-2}\,\mathrm{W}\left(X(s),X(t)\right)^2 \leq \sup_{s\leq h\leq t}\int(\pi Hp(.,h)(x)-2^{-1}x)^2\,p(x,h)\,dx$$

Proof: Let. C be a variable with Cauchy distribution and free w.r.t. $\{X,S\}$. Let .If $g\in C^{\infty}(\mathbf{R})$ be such that $\|g'\|_{\infty}\leq 1, g(x)=x$ if $|x|\leq\|X\|+1$ and $g'(x)=0$ if $|x|\geq\|X\|+2$. We apply $X+\lambda\mathrm{C}$ in place of X. Let.

$$Z(t,\lambda)=e^{-t/2}\left(X+\lambda C\right)+\left(1-e^{-t}\right)^{1/2}S=X(t)+e^{-t/2}\lambda C$$

Then $\mathrm{g}(\mathrm{Z}(\mathrm{t},\lambda))$ is an operator of $\text{norm}\leq\|\mathrm{X}\|+2$ and converges in distribution to $X(t)$, Moreover the distribution of $Z(t,\lambda)$ is given by the density $p_{e^{-t/2}\lambda}*p(.,t)$ and will he denoted by $p(.,t,\lambda)$. In view of the L^{m}-continuity of $p(.,t)(1<m<\infty)$ [47] it is easy to see that

$$\limsup_{\lambda\downarrow 0}\left(\sup_{s\leq h\leq t}\int_{\operatorname{supp} g'}\left(\pi Hp(.,h)(x)-2^{-1}x\right)^2 p(x,h,\lambda)dx\right)\leq \sup_{s\leq h\leq t}\int\left(\pi Hp(.,h)(x)-2^{-1}x\right)^2 p(x,h)dx$$

From now on we return to the context of bounded variables X. If the distribution of X is Lebesgue absolutely continuous and has density p which is L^3. then $\frac{1}{2}\mathrm{J}\ (X)=\pi Hp(X)$ where $\mathrm{J}\ (X)$ is the conjugate variable (a.k.a. free Brownian gradient, a.k.a. non-commutative Hilbert transform) (see [44,52]) and $\Phi(X)=\tau\left(\mathrm{J}\ (X)^2\right)=4\pi^2\int\left(Hp(x)\right)^2 p(x)dx=\frac{4}{3}\pi^2\int p^3(x)dx$

is the free Fisher information (see [45,44] up to different normalizations). The quantity occurring in Lemma (2.5):

$$I(X)=4\int\left(\pi Hp(x)-2^{-1}x\right)^2 p(x)dx=\tau\left(\left(\mathrm{J}\ (X)-X\right)^2\right)=\Phi(X)-2+\tau\left(X^2\right);$$

is a generalization of the free Fisher information for Ornstein-Uhlenbeck processes (see [44]). The inequality in Lemma (2.5): can also be written

$$4(t-s)^{-2}\,\mathrm{W}\left(X(s),X(t)\right)^{2}\leq\sup_{s\leq h\leq t}I\left(X(h)\right). \tag{14}$$

The free entropy. The free entropy of X with distribution $\mu=\mu_X$ is

$$\chi(X)=\iint\log|s-t|\,d\mu(s)d\mu(t)+\frac{3}{4}+\frac{1}{2}\log(2\pi)$$

(see [45,44] up to different constants) and we have

$$\chi(\alpha X)=\chi(X)+\log|\alpha|\ \text{ and }\ \lim_{\varepsilon\downarrow 0}\varepsilon^{-1}\left(\chi\left(X+\varepsilon^{1/2}S\right)-\chi(X)\right)=2^{-1}\Phi(X)$$

The quantity we shall use in estimating the distance to the semicircle distribution is a modified free entropy adapted to the free Ornstein-Uhlenbeck process ([202]):

$$\dot{\sum}(X)=-\chi(X)+\chi(S)+\frac{1}{2}\tau\left(X^{2}\right)-\frac{1}{2}=\frac{1}{2}\tau\left(X^{2}\right)-\iint d\mu(s)\,d\mu(t)\log|s-t|-\frac{3}{4}$$

We have

$\lim_{t\to\infty}\dot{\sum}\left(X(t)\right)=0$ and

$$\frac{d}{dt}\dot{\sum}\left(X(t)\right)=\frac{d}{dt}\left(\frac{1}{2}-\chi\left(X+\left(e^{t}-1\right)^{\frac{1}{2}}S\right)+\frac{1}{2}e^{-t}\tau\left(X^{2}\right)+\frac{1}{2}\left(1-e^{-t}\right)\right)$$

$$=2^{-1}\left(1-e^{-t}\Phi\left(X+\left(e^{t}-1\right)^{\frac{1}{2}}S\right)-e^{-t}\tau\left(X^{2}\right)+e^{-t}\right)=2^{-1}\left(1-\Phi\left(X(t)\right)-\tau\left(X(t)^{2}\right)+1\right)=-2^{-1}I(X).$$

Note also that in [202] using the logarithmic Sobolev inequality for X (see [44]) , it is shown that

$$\dot{\sum}\left(X(t)\right)\leq 2^{-1}I\left(X(t)\right), \tag{15}$$

which is a logarithmic Sobolev inequality for the Ornstein-Uhlenbeck process.

Lemma (2.6)[198]: Assume X,Y are bounded and self-adjoint, then if $t>0$ we have

$$\limsup_{\varepsilon\to 0}|\varepsilon|^{-1}\left|\mathrm{W}\left(Y,X(t+\varepsilon)\right)-\mathrm{W}\left(Y,X(t)\right)\right|\leq 2^{-1}\left(I\left(X(t)\right)\right)^{\frac{1}{2}}$$

We now have all ingredients to get an estimate for $\mathrm{W}(X,S)$ which is similar in the free context to an' inequality of Talagrand in the classical setting (

[69],[179,224])

Theorem (2.7)[198]: $W(X,S)^2 \le 2\dot{\sum}(X)$**.**

Proof: Because of the semicircular maximum for χ we have

$\chi(X) \le \chi(S) + 2^{-1}\log(\tau(X^2))$ so that $\dot{\sum}(X) \ge 2^{-1}(\tau(X^2) - (1+\log\tau(X^2))) \le 0$.

Thus it will suffice to prove that $W(X,S) - \left(2\dot{\sum}(X)\right)^{1/2} \le 0$. By Lemma **(2.7)** the inequality (15) and the formula for the derivative of $\dot{\sum}(X(t))$, we have for $t > 0$,

$$\liminf_{\varepsilon\to 0} \varepsilon^{-1}\left(W(X(t+\varepsilon),S) - \left(2\dot{\sum}(X(t+\varepsilon))\right)^{\frac{1}{2}} - W(X(t),S) + \left(2\dot{\sum}(X(t))\right)^{\frac{1}{2}} \right)$$

$$\ge -2^{-1}(I(X(t)))^{1/2} + 2^{-1/2} I(X(t))(2\dot{\sum}(X(t)))^{-1/2} \ge -2^{-1}(I(X(t)))^{1/2} + 2^{-1/2+1/2} I(X(t))((IX(t)))^{-1/2} = 0$$

Hence $W(X(t),S) - \left(2\dot{\sum}(X(t))\right)^{1/2}$ is an increasing function and we have

$$\lim_{t\to\infty}\left(W(X(t),S) - \left(2\dot{\sum}(X(t))\right)^{\frac{1}{2}} \right) = 0.$$

because of the semicircular maximum and lower semicontinuity of χ. It follows that. .

$$W(X(t),S) - \left(2\dot{\sum}(X(t))\right)^{1/2} \le 0$$

if $t > 0$. To get the inequality for $t = 0$, remark that $X(t)$ is norm-continuous so that $W(X(t),S)$ tends to $W(X,S)$ as $t \to 0$. On the other hand. by lower semicontinuity of χ ,

$$\liminf_{t\downarrow 0} -\left(\left(-\dot{\sum}(X(t+\varepsilon)) \right)^{\frac{1}{2}} \right) \ge -\left(\dot{\sum}(X)\right)^{1/2}$$

Because of the coincidence of the free and classical Wasserstem distance for single self-adjoint variables, the preceding theorem can also be written in terms of probability measures for the classical distance. Let μ be a compactly supported probability measure on **R** and σ a $(0,1)$ semicircle distribution. Then we have

$$(\mathrm{W}(\mu,\sigma))^2 \leq \int x^2 d\mu(x) - 2\int\int d\mu(s)\, d\mu(t) \log|s-t| - \frac{3}{2}.$$

Assume A is a unital C* algebra $l \in B \subset A$ a C^*-subalegbra and A is generated by $BU\{a_1,...,a_n\}$. If $\theta \in TS(B)$ let $TS(A:B,\theta) = \{\tau \in TS(A) | \tau | \mathrm{B} = \theta\}$. If $\tau_j \in TS(A_j : B,\theta)$, where $1 \in B \subset A_{j,}$ j = 1,2, let $TS(A_{1*B}.A_2{:}B,\theta;\tau_1,\tau_2) = \{\tau \in TS(A_{1*B}A_2 : B,\theta) | \tau \circ \sigma_j = \tau_j, j = 1,2\}$ where $A_{l\,*B}A_2$ is the full free product with amalgamation over B. The relative Wasserstein metric is then

$$\mathrm{W}_p(\tau_1,\tau_2{:}\theta) = \inf\{\left\| (\left|\sigma_1(\mathrm{a}_j) - a_2(a_j)\right|_{p,\tau}\ 1 \leq j \leq n \right\|_p | \tau \in TS(A_{*B}A : B,\theta;\tau_1,\tau_2)\},$$

Where $(\tau_j \in TS(A:B,\theta)$, $\mathrm{j} = 1,2.$

Proposition (2.8) [198]: (a) $\mathrm{W}_p(.,.\theta)$ is a metric on $TS(A:B,\theta)$. (b) The infimum in the definition of $\mathrm{W}_\mathrm{p}(\tau_1,\tau_2:\theta)$ is attained. (c) Let $\tau_1^{(k)},\tau_1,\tau_2^{(k)},\tau_2 \in T(\mathrm{A}:\mathrm{B},\theta)$ and assume $\tau_j^{(k)}$ converges weakly to τ_j as $k \to \infty (j=1,2)$. Then

$$\liminf_{k\to\infty} \mathrm{W}_p\left(\tau_1^{(k)},\tau_2^{(k)}:\theta\right) \geq \mathrm{W}_p(\tau_1,\tau_2:\theta).$$

There is also a corresponding version of the relative metric for $n-$ tuples of noncommutative random variables. Let $(\mathrm{M}_\mathrm{i},\tau_1),(\mathrm{M}_2,\tau_2)$ be tracial W^*-probability space, so that $1 \in B \subset M_j, \tau_j | B = \theta, j = 1,2.$. If $X_1,...,X_\mathrm{n} \in M_1$ and $\mathrm{Y}_1,...,\mathrm{Y}_\mathrm{n} \in M_2$ we define $\mathrm{W}_\mathrm{p}((X_1,...,X_\mathrm{n}),(Y_1,...,Y_\mathrm{n}) : \mathrm{B})$ as the infimum of $\left\| \left(|X'_j - Y'_j|_p \right)_{1\leq j\leq n} \right\|_p$ over $2n-$tuples $(X'_1,...,X'_n,Y'_1,...,Y'_n)$ in traciai W^*-probability spaces (M_3,τ_3) so that $I \in B \subset M_3, \tau_3 | B = \theta$ and $(X_1,...,X_\mathrm{n},B),(X'_1,...,X'_\mathrm{n},\mathrm{B})$ and respectively $(Y_1,...,Y_n,B)(Y'_1,...,Y'_n,B)$ have the same *- distributions.

.Note also that in case (B,θ) is given by generators $(\mathrm{Z}_1,..., \mathrm{Z}_\mathrm{p})$ this leads to relative metrics denoted

$$\mathrm{W}_\mathrm{p}(X_1,...,X_\mathrm{n}),(\mathrm{Y}_1,...,\mathrm{Y}_\mathrm{n}) : (Z_1,...,\mathrm{Z}_\mathrm{p}))\ .$$

There are many higher dimensional metric quantities which appear naturally in this context of optimization problems on trace-state spaces. Here we only want to give a few examples, to bring the reader's attention to this unexplored structure of trace-state space. The idea is quite simple. Given an m-tuple $(\tau_1,...,\tau_m)$ of trace-states, then an

element $\tau \in TS(A_{*}..._{*}A;\tau_1,...,\tau_m)$ yields a m-tuple of vectors

$$a_k(\tau)=(\sigma_k(a_j))_{1\le j\le n} \in (L^2(A_{*}..._{*}A;\tau_1,...,\tau_m))^n,$$

Where $1\le k\le m$.

Then if w is some geometric. quantity associated with an n-tuple of points in a Hilbert space, we may consider $\mathrm{w}(\mathrm{a}_1(\tau),...,\mathrm{a}_\mathrm{n}(\tau))$ and then define.

$$\mathrm{W}(\tau_1,...,\tau_m)=\inf\{\phi(a_1(\tau),...,a_n(\tau))|\tau\in TS(\mathrm{A}\ _{*}..._{*}\mathrm{A};\tau_1,...,\tau_\mathrm{m})\}$$

Two examples of such w are $volp(\mathrm{h}_1,...,\mathrm{h}_\mathrm{m})$ and $volS(\mathrm{h}_1,...,\mathrm{h}_\mathrm{m})$, the m- -dimensional volume of the parallelipiped defined by the vectors $\mathrm{h}_1,...,\mathrm{h}_\mathrm{m}$ and respectively. The $(m-1)$ -dimensional volume of the simplex with vertices $\mathrm{h}_1,...,\mathrm{h}_\mathrm{m}$.In case A is commutative there are corresponding "classical" quantities $W_{ab}(\tau_1,...,\tau_n)$,where the supremum is over $TS_{ab}(A_{*}..._{*}A;\tau_1,...,\tau_m)=TS(A_{*}..._{*}A;\tau_1,...,\tau_m\bigcap T(A\otimes...\otimes A)$ where $A^{\otimes n}$ is viewed as a quotient of A^{*n}. We conclude with a few remarks about the volumes of parallelepipeds $volp$ If H is aHilbert space and $\mathrm{h}_1,...,h_m\in\mathrm{H}$ let

$$h_i\wedge...\wedge h_m=(m!)^{-1/2}\sum_{\gamma\in\varphi_m}\varepsilon(\gamma)h_{\gamma(1)}\oplus...\otimes h_{\gamma(m)}\in\mathrm{H}^{\otimes m}$$

be the exterior product, where ϕ_m is the permutation group and $\varepsilon(\gamma)$ the sign of the permutation. The norm $\|h_1\wedge...\wedge h_m\|$ is the norm from $\mathrm{H}^{\otimes m}$ and by definition $volp\ (h_1,...,h_m)=\|h_1\wedge...\wedge h_m\|$.

Proposition (2.9)[198]: (i) $volp(\tau_1,...,t_q).volp(\tau_{q+1},...\tau_{q+r})\ge volp(\tau_1,...,\tau_{q+\tau})$;

(ii) $vol\ p(\tau_1)=\left(\sum_{1\le k\le n}\tau_1(a_k^*a_k)\right)^{1/2}$;(iii) If $a_k\ge 0(1\le k\le n)$ then $2^{-1}\mathrm{volp}(\tau_1,\tau_2)$ is the area of a triangle with sides $volp(\tau_1), volp(\tau_2)$, $\mathrm{W}(\tau_1,\tau_2)$;(iv) The jnfimum in the definition of $volp$ is attained;(v) Let $\tau_j^{(k)},\tau_j\in TS(A), 1\le j\le m, k\le N$, and assume $\tau_j^{(k)}$ converges weakly to τ_j as $k\to\infty$. Then $\liminf_{k\to\infty} volp(\tau_1^{(k)},...,\tau_m^{(k)})\ge volp(\tau_1,...,\tau_m)$.

Like the proof of the triangle inequality for W, the proof of (i) is based on free products $\tau' {}_* \tau''$ where $\tau' \in TS\left(A^{*q};\tau_1,...,\tau_q\right), \tau'' \in TS\left(A^{*r};\tau_{q+1,...,}\tau_{q+r}\right)$. Also (iv) and (v) have quite similar proofs to corresponding properties of W. The condition $a_k \geq 0, (1 \leq \mathrm{k} \leq \mathrm{n})$ in (iii) insures that the angle between the vectors $a_1(\tau)$ and $a_2(\tau)$ is acute and under such a condition it is clear that the area of the triangle with sides $a_1(\tau)$, $a_2(\tau)$ of constant length is minimum at the same time with the third side $\left|a_1(\tau) - a_2(\tau)\right|_2$.

Chapter 3

Free Brunn-Minkowski and Talagrand Inequalities

The method is used to extend to the free setting the Otto-Villani theorem starting that the logarithmic Sobolev inequality implies the transportation cost inequality. We also discuss the convergence, fluctuations and large deviations of the energy of the eigenvalues of β ensembles ,which ,as an application of Talagrand inequality gives in particular yet another proof the convergence of the eigenvalue distribution to the semicircle law.

Section (3.1): One dimensional Brunn-Minkowski inequality

In its functional form, the Brunn-Minkowski inequality indicates that whenever $\theta \in (0,1)$ and u_1, u_2, u_3 are non-negative measurable functions on $\mathbf{R}^n$ such that

$$u_3\left(\theta x + (1-\theta)y\right) \geq u_1(x)^{\theta} u_2(y)^{1-\theta} \text{ for all } x, y \in \mathbf{R}^n,$$

then

$$\int u_3 dx \geq \left(\int u_1 dx\right)^{\theta} \left(\int u_2 dx\right)^{1-\theta}$$

The Brunn-Minkowski inequality has been used recently in the investigation of functional inequalities for strictly log-concave densities such as logarithmic Sobolev or transportation cost inequalities (cf. [174, 23], [79] (cf. [65]). Given a continuous function $Q:\mathbf{R}\to\mathbf{R}$ such that $\lim_{|x|\to\infty} |x| e^{-\varepsilon Q(x)} = 0$ for every $\varepsilon > 0$, set

$$\tilde{Z}_N(Q) = \int \Delta_N(x)^2 e^{-N\sum_{k=1}^{N} Q(x_k)} dx$$

where $A = \{x_1 < x_2 < \ldots < x_N\} \subset \mathbf{R}^N$ and $\Delta_N(x) = \prod_{1\leq k<\ell\leq N}(x_\ell - x_k)$ is the Vandermonde determinant. The large deviation theorem of [45] and [79] (see also [132]) indicates that.

$$\lim_{N\to\infty} \frac{1}{N^2} \log \tilde{Z}_N(Q) = \varepsilon_Q(\mathrm{V}Q) \tag{1}$$

where, for every probability measure V on $\mathbf{R}$,

$$\sum_Q(\mathrm{V}) = \iint \log|x-y| d\mathrm{V}(x) d\mathrm{V}(y) - \int Q(x) d\mathrm{V}(x).$$

is the weighted energy integral with external (compactly supported) measure V_Q maximizing $\sum_Q$ (cf. [52,65]). (For the choice of $Q(x) = \frac{x^2}{2}, \mathrm{V}_Q$ is the semicircle law.)

Let u_1, u_2, u_3 be real-valued continuous functions on $\mathbf{R}$ such that, for every $\varepsilon > 0$, $\lim_{|x|\to\infty} |x| e^{-\varepsilon U_i(x)} = 0, i = 1,2,3$. Set

$$u_i(x) = \Delta_N(x)^2 e^{-N\sum_{k=1}^N U_i(x_k)} \mathrm{I}_A(x), \quad x \in \mathbf{R}^N, i = 1,2,3.$$

Since $-\log \Delta_N$ is convex on the convex set A, assuming that, for some $\theta \in (0,1)$ and all $x, y \in \mathbf{R}$, $U_3(\theta x + (1-\theta) y) \le \theta U_1(x) + (1-\theta) U_2(y)$, the Brunn-Minkowski theorem applies to u_1, u_2, u_3 on $\mathbf{R}^N$ to yield

$$\tilde{Z}_N(U_3) \ge \tilde{Z}_N(U_1)^\theta \tilde{Z}_N(U_2)^{1-\theta}$$

Taking the limit (1) immediately yields the following free analogue of the functional Brunn-Minkowski inequality on $\mathbf{R}$.

Theorem (3.1.1) [171]: Let u_1, u_2, u_3 be real-valued continuous functions on $\mathbf{R}$ such that, for every $\varepsilon > 0$, $\lim_{|x|\to\infty} |x| e^{-\varepsilon U_i(x)} = 0, \quad i = 1,2,3$. Assume that for some $\theta \in (0,1)$ and all $x, y \in \mathbf{R}$

$$u_3(\theta x + (1-\theta) y) \le u_1(x) + (1-\theta) u_2(y)$$

Then

$$\sum_{U_3}(\mathrm{V}_{U_3}) \ge \theta \Sigma_{U_1}(\mathrm{V}_{U_1}) + (1-\theta) \sum_{U_2}(\mathrm{V}_{U_2})$$

The free analogue of Shannon's entropy power inequality due to Szarek and Voiculescu [250] may be recovered along the same lines.

We next show how the preceding free Bnmn-Minkowski inequality may be used, following the classical case, to recapture both the free logarithmic Sobolev inequality of Voiculescu [44] (in the form put forward in [197] and extended in [200]) and the free quadratic transportation cost inequality of [198,66] for quadratic and more general strictly convex potentials Q.

Let Q be a real-valued continuous function on $\mathbf{R}$ such that $\lim_{|x|\to\infty} |x| e^{-\varepsilon Q(x)} = 0$ for every $\varepsilon > 0$. For V, probability measure on $\mathbf{R}$, define the free entropy of V (with respect to V_Q) [44, 200,202] as

$$\dot{\sum}\left(\mathrm{V} \middle| \mathrm{V}_Q\right) = \sum\nolimits_Q\left(\mathrm{V}_Q\right) - \sum\nolimits_Q(\mathrm{V}) \geq 0$$

If $\phi : \mathbf{R} \to \mathbf{R}$ is bounded and continuous, it is convenient to set below

$\lambda_Q(\phi) = \sum_{Q-\varphi}\left(\mathrm{V}_{Q-\varphi}\right) - \sum_Q\left(\mathrm{V}_Q\right)$. For every probability measure V on $\mathbf{R}$,

$$\lambda_Q(\phi) \geq \int \phi d\,\mathrm{V} + \sum\nolimits_Q(\mathrm{V}) - \sum\nolimits_Q\left(\mathrm{V}_Q\right) = \int \phi\, d\,\mathrm{V} - \dot{\sum}\left(\mathrm{V} \middle| \mathrm{V}_Q\right)$$

with equality for $\mathrm{V} = \mathrm{V}_{Q-\varphi}$. In particular $\lambda_Q(\phi) \geq \int \phi d\mathrm{V}_Q$.

Assume now that (Q is C^1 and such that) $Q(x) - c/2x^2$ is convex for some $C > 0$. For bounded continuous functions $f, g : \mathbf{R} \to \mathbf{R}$ such that $g(x) \leq f(y) + c/2|x-y|^2$, we may apply the free Bnmn-Minkowski theorem, as in the classical case (cf. [173]), to $U_1 = Q - (1-\theta)g, U_2 = Q + \theta f$ and $U_3 = Q$.

Thus, by the theorem, $\lambda_Q((1-\theta)g) + \frac{1-\theta}{\theta}\lambda_Q(-\theta f) \leq 0$. As $\theta \to 0$, it follows that for every probability measure V,

$$\int g d\mathrm{V} - \int f d\mathrm{V}_Q \leq \dot{\sum}\left(\mathrm{V} \middle| \mathrm{V}_Q\right)$$

(in other words $\lambda_Q(g) \leq \int f d\mathrm{V}_Q$). By the Monge-Kantorovitch-Rubinstein theorem (see [23]), this is the dual form of the free quadratic transportation cost inequality.

$$\mathrm{W}_2\left(\mathrm{V}, \mathrm{V}_Q\right)^2 < \frac{1}{c}\dot{\sum}\left(\mathrm{V} \middle| \mathrm{V}_Q\right) \qquad (2)$$

recently put forward in [198] for the semicircle law associated to the quadratic potential, and in [66] for strictly convex potentials (where $\mathrm{W}_2\left(\mathrm{V}, \mathrm{V}_Q\right)$ is the Wasserstein distance between V and V_Q).

The free logarithmic Sobolev inequality of [44], extended to strictly convex potentials in [200], follows in the same way from the free Brunn-Minkowski theorem. We follow [200] where the matrix approximation is used similarly to this task. Fix a probability measure V with compact support and smooth density p on $\mathbf{R}$. Define a C^1 function R on $\mathbf{R}$ such that $P(x)=2\int\log|x-y|d\nu(y)$ on $\operatorname{supp}(\nu)$, $P(x)=Q(x)$ for $|x|$ large, and such that $P(x)\geq 2\int\log|x-y|d\nu(y)$ every-where. By the uniqueness theorem of extremal measures of weighted potentials (cf.[52]), it is easily seen that the energy functional $\Sigma_{\mathbf{R}}$ is maximized at the unique point $\nu_{\mathbf{R}}=\nu$. Define then f. with compact support, by $f=Q-\mathbf{R}+C$ where the constant $C\left(=\Sigma_Q(\nu_Q)-\Sigma_{\mathbf{R}}(\nu_{\mathbf{R}})\right)$ is chosen so that $\lambda_Q(f)=0$.

Let $g_t(x)=\inf_{y\in\mathbf{R}}\left[f(y)+\frac{1}{2t}(x-y)^2\right],\ t>0,\quad x\in\mathbf{R}$, be the infimum-convolution of f with the quadratic cost, solution of the Hamilton-Jacobi equation $\partial_t g_t+\frac{1}{2}g_t'^2=0$ with initial

condition f. As in the classical case (cf. [173]), apply the Brumn-Minkowski theorem to $U_1=Q-\frac{1}{\theta}g_t,\ t=\frac{1-\theta}{c^\theta},\ U_2=Q,\ U_3=Q-f$ to get that $j_Q\left((1+ct)g_t\right)\leq j_Q(f)=0$ for every $t>0$. In particular therefore,

$\int(1+ct)g_t d\nu\leq\tilde{\Sigma}\left(\nu|\nu_Q\right)$,and, since $\nu=\nu_{Q-f}$, as $t\to 0$,

$$\tilde{\Sigma}\left(\nu|\nu_Q\right)=\int f d\nu\leq\frac{1}{2c}\int f'^2 d\nu$$

Now $f'=Q'-Hp$ where $Hp(x)=\text{p.v.}\int\frac{2p(y)}{x-y}$ is the Hilbert transform (up to a multiplicative factor) of the (smooth) density p of ν. Hence the preceding atitounts to the free logarithmic Sobolev inequality

$$\tilde{\Sigma}\left(\nu|\nu_Q\right)\leq\frac{1}{2c}\int[Hp-Q']^2 d\nu=\frac{1}{2c}I\left(\nu|\nu_Q\right)\qquad(3)$$

as established in [200], where $I\left(\nu|\nu_Q\right)$ is known as the free Fisher information of ν with respect to ν_Q [44, 197]. Careful approximation arguments to reach arbitrary probability measures ν (with density in $L^3(\mathbf{R})$ are described in [67].

The Hamilton-Jacobi approach may be used to prove, as in the classical case, the free analogue of the OttoVillani theorem [69] (cf. [23, 256, 173]) stating that, for a given probability measure $d\mu=e^{-Q}dx$ on $\mathbf{R}$ (with a C^1 potential Q such that $\lim_{|x|\to\infty}|x|e^{-\varepsilon Q(x)}=0$ for every $\varepsilon>0$), the free logarithmic Sobolev inequality (3) always implies the free transportation cost inequality

(2). To this task, given a compactly supported C^1 function f on $\mathbf{R}$ and $a \in \mathbf{R}$, set $j_t = j_Q((a+ct)g_t)$ and $f_t = (a+ct)g_t - j_t$ so that $j_Q(f_t) = 0$. Denote for simplicity by V_t the extremal measure for the potential $Q - f_t$. Then the logarithmic Sobolev inequality (3) can be expressed as $\int f_t dV_t \le \frac{1}{2c}\int f_t'^2 dV_t$. In other words,

$$c(a+ct)\int g_t dV_t - cl_t \le -(a+ct)^2 \int \partial_t g_t dV_t$$

On the support of V_t (cf. [52]),

$$2\int \log|x-y| dV_t(y) = Q - f_t + C_t$$

where $C_t = \int\int \log|x-y| dV_t dV_t + \Sigma_{Q-f_t}(V_t)$ Since $l_Q(f_t) = \Sigma_{Q-f_t}(V_t) - \Sigma_Q(V_Q) = 0$,it follows

that $\int \partial_t f_t dV_t = 0$. Therefore, $cj_t \ge (a+ct)\partial_t j_t$ and hence $(a+ct)^{-1} j_t$ is non-increasing in t. In particular $\frac{1}{a+1} j_{1/c} \le \frac{1}{a} j_0$, which for a = 0 amounts to $j_Q(g) \le \int f dV_Q$, that is the dual form of (2). This approach through the Hamilton-Jacobi equations has some similarities with the use of the (complex) Burgers equation in [198].

Section (3.2) Semicircular Law and Energy of the eigenvalues of Beta Ensembles

In [179] Talagrand proves the transportation-cost inequality to the Gaussian measure. The one dimensional version for the Gaussian measure $\gamma(dx) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$

reads as

$$\left(\mathrm{W}(\mu,\gamma)\right)^2 \le 2H(\mathrm{V}|\gamma), \qquad (4)$$

where $\mathrm{W}_2(\mu,\gamma)$ 1is the Wasserstien distance defined below by (8) and the relative entropy is

$$H(\mathrm{V}|\gamma) = \begin{cases} \int f(x)\log f(x)\, d\gamma(x) & \text{if } \mathrm{V}(dx) = f(x)\gamma(dx) \\ \infty & \text{if } \mathrm{V} \text{ is singular to } \gamma \end{cases}$$

In the context of free probability, Biane and Voiculescu provided in [198] a free version of this:

$$\left(W_2(\mu,\sigma)\right)^2 \le 2\left(E(\mu)-E(\sigma)\right) \tag{5}$$

where $E(\mu)=\frac{1}{2}\int x^2\mu(dx)-\int\int\log(|x-y|)\mu(dx)\mu(dy)$ is the free energy of μ and $\sigma(dr)=\frac{1}{2\pi}I_{[-2,2]}(x)\sqrt{4-x^2}dx$ is the semicircular law, the minimizer of $E(\mu)$ over all probability measures on the real line. The role of the relative entropy is played here by the difference of the free energy of μ and the semicircular.

Using random matrix approximations, Hiai, Petz and Ueda proved in [6] the following extension of (5),

$$\rho\left(W_2(\mu,\mu_Q)\right)^2 \le E^Q(\mu)-E^Q(\mu_Q) \tag{6}$$

where $\rho>0$ and $Q:\mathbf{R}\to\mathbf{R}$ is a function so that $Q(x)-\rho x^2$, is convex and

$$E^Q(\mu)=\int Q(x)\mu(dx)-\int\int\log|x-y|\mu(dx)\mu(dy)$$

Here μ_Q is the minimizer of E^Q on the set of all probability measures on the real line. They also prove a version of this for measures supported on the circle T :

$$(\rho+1/4)\left(W_2(V,V_Q)\right)^2 \le E^Q(V)-E^Q(V_Q) \tag{7}$$

where $Q:T\to\mathbf{R}$ so that $Q(e^{ix})-\rho x^2$ is convex on $\mathbf{R}, \rho>-1/4$ and μ_Q is the minimizer of the functional E^Q on probability measures on the unit circle T .

Another proof of (5) is given in [171] via a Brunn-Minkovsky inequality for free probability.

The following result is an obvious one but is the key to our problem.

Lemma (3.2.1) [102]: Let $f:[0,1]\to\mathbf{R}$ be a convex function with the property that $f(0)=0$ and there exists $a\ge 0$ so that

$$f(t)\ge -at^2 \quad for \quad t\in[0,1]$$

Then

$$f(t)\ge 0 \quad for\, all \quad t\in[0,1]$$

Proof: It follows from the assumptions that for any $\epsilon > 0$, if $\delta_\epsilon = \min(1, \epsilon / a)$, then $f(t) \geq -t\epsilon \; for \; t \in [0, \delta_\epsilon]$. Now, since f is convex, one gets $f(mt) \geq mf(t) \geq -mt\epsilon$ for any integer m with $mt \leq 1$, and therefore, $f(t) \geq -\epsilon t$ for any $t \in [0,1]$. Since this is true for any $\epsilon > 0$, we get $f(t) \geq 0$ for any $t \in [0,1]$. □

In the following, $p(\Omega)$ denotes the set of all probability measures on Ω, and for two probability measures with finite second moment on $p(\mathbf{R})$ or $p(\mathrm{T})$, where $\mathrm{T} = \{z \in \mathrm{C} : |z| = 1\}$, we define $\mathrm{W}_2(\mu, \mathrm{V})$, the Wasserstein distance by

$$\mathrm{W}_2(\mu, \mathrm{V}) = \sqrt{\inf_{\pi \in \prod(\mu, v)} \int\int |x - y|^2 \, d\pi(x, y)} \tag{8}$$

Here $\prod(\mu, \mathrm{V})$ is the set of probability measures on $\mathbf{R}^2$ with marginal distributions μ and v, and it can be shown that there is at least one solution $\pi \in \prod(\mu, \mathrm{V})$ to this minimization problem.

If μ and V are two measures on $\mathbf{R}$ with F and G their cumulative distribution functions (i.e. $F(x) = \mu((-\infty, x])$), then in [23] states that

$$\left(\mathrm{W}_2(\mu, \mathrm{V})\right)^2 = \int_0^1 \left|F^{-1}(t) - G^{-1}(t)\right|^2 dt \tag{9}$$

where F^{-1} denotes the generalized inverse of F.

Theorem (3.2.2) [102]: Let $Q : \mathbf{R} \to \mathbf{R}$ be a function so that $Q(x) - \rho x^2$ is convex for a certain $\rho > 0$. If μ_Q is a solution to the minimization problem

$$I^Q = \inf_{\mu \in \mathrm{P}(\mathbf{R})} \mathrm{E}^Q(\mu) \tag{10}$$

where

$$\mathrm{E}^Q(\mu) = \int Q(x)\mu(dx) - \int\int \log|x - y| \, \mu(dx)\mu(dy) \tag{11}$$

then for any $\mu \in p(\mathbf{R})$, we have

$$\rho\left(\mathrm{W}_2(\mu, \mu_Q)\right)^2 \leq \mathrm{E}^Q(\mu) - I^Q \tag{12}$$

In particular, the minimization problem (10) has a unique solution.

Proof: There exist constants C_1 and C_2 so that

$$Q(x)-\rho x^2 \geq C_1 \text{ and } \log(|x-y|) \geq -\frac{\rho}{4}(x^2+y^2)+c_2.$$

Then for a certain C, we get that

$$\frac{1}{2}(Q(x)+Q(y))-\log(|x-y|) \geq \frac{\rho}{4}(x^2+y^2)+C \geq C \tag{13}$$

and this in turn implies that the infimum in (10) is finite (since $E^Q(\mu)$ is finite for μ the uniform distribution on [0,1]) and in particular $\int Q(x)\,d\mu_Q(x)$, and $\int\int \log|x-y|\,d\mu_Q(x)d\mu_Q(y)$ are finite, which means that μ_Q has finite second moment and no atoms.

Since $E^Q(\mu) > -\infty$, we may assume that $\mathrm{E}^Q(\mu)$ is finite, otherwise there is nothing to prove. Then, $\int\int \log|x-y|\,\mu(dx)\mu(dy)$ and $\int Q(x)\,\mu(dx)$ are finite. In particular, μ has finite second moment and no atoms.

Taking F_μ and $F_{\mu Q}$ the cumulative distributions of μ, μ_Q and F^{-1}, F_Q^{-1} their generalized inverses, set $\theta(x) = F^{-1}(F_Q(x))$. According to [23] and the discussion following thereafter, the minimizing measure π from (8) is the distribution of $x \to (x, \theta(x))$ under μ_Q. In this case, the inequality we want to prove becomes

$$\rho\int\int |x-\theta(x)|^2\,\mu_Q(dx) \leq \int Q(x)\,\mu(dx) - \int\int \log|x-y|\,\mu(dx)\mu(dy) - I^Q$$

Let $f:[0,1] \to \mathbf{R}$ be given by

$$f(t) = -\rho t^2 \int |\theta(x)-x|^2\,\mu_Q(dx) + \int Q(t\theta(x)+(1-t)x)\,\mu_Q(dx)$$
$$-\int\int \log(|t(\theta(x)-\theta(y))+(1-t)(x-y)|)\mu_Q(dx)\,\mu_Q(dy) - I^Q$$

Notice here that f is well defined. Indeed, Q is convex, hence bounded below and because $\int Q(\theta(x))\,\mu_Q(dx) = \int Q(x)\,\mu(dx)$ and $\int Q(x)\,\mu_Q(dx)$ are both finite, one concludes that $\int Q(t\theta(x)+(1-t)x)\,\mu_Q(dx)$ is finite too. One the other hand, there is a $C>0$ so that for any $t \in [0,1]$,

$$-\log(|t(\theta(x)-\theta(y))+(1-t)(x-y)|) \geq -C(\theta(x)^2+\theta(y)^2+x^2+y^2)-C,$$

which, combined with the finiteness of the second moment of μ and μ_Q, results with (for a constant C)

$$-\int\int \log\left(\left|t\left(\theta(x)-\theta(y)\right)+(1-t)(x-y)\right|\right)\mu_Q(dx)\,\mu_Q(dy) > C \text{ for all } t\in[0,1].$$

Now, since θ is a no decreasing function we can write

$$-\int\int \log\left(\left|t\left(\theta(x)-\theta(y)\right)+(1-t)(x-y)\right|\right)\mu_Q(dx)\,\mu_Q(dy)$$
$$=-2\int\int_{x>y} \log\left(t\left(\theta(x)-\theta(y)\right)+(1-t)(x-y)\right)\mu_Q(dx)\,\mu_Q(dy),$$

which combined with the convexity of $-\log$ on $(0,\infty)$ and the finiteness of $\int\int \log|x-y|\mu_Q(dx)\,\mu_Q(dy)$ and $\int\int \log|x-y|\mu(dx)\,\mu(dy)$, yields the fact that

$$t\to -\int\int \log\left(\left|t\left(\theta(x)-\theta(y)\right)+(1-t)(x-y)\right|\right)\mu_Q(dx)\,\mu_Q(dy) \tag{14}$$

is well defined and convex.

The inequality (12) is now equivalent to $f(I)\geq 0$. To show this, we apply Lemma (3.2.1): The convexity follows easily from the convexity of $Q(x)-\rho x^2$ and (14). Now if ν_t is the distribution of $x\to t\theta(x)+(1-t)x$ under μ_Q, then the minimization property of μ_Q implies that

$$f(t)\geq -\rho t^2\int\int \left|\theta(x)-x\right|^2\mu_Q(dx), \text{ for } t\in[0,1]$$

and then, Lemma (3.2.1) : shows that $f(I)\geq 0$ for any $t\in[0,1]$.

The existence statement follows from the lower continuity of E^Q. For a proof of the existence and compactness of the support of μ_Q (see [206]).

Corollary (3.2.4) [102]: Let $\sigma(dx)=\frac{1}{2\pi}\mathrm{I}_{[-2,2]}(x)\sqrt{4-x^2}\,dx$ be the semicircular law on $[-2,2]$. Then for any $\mu\in p(\mathbf{R})$,

$$\frac{1}{2}\left(\mathrm{W}_2(\mu,\sigma)\right)^2\leq \frac{1}{2}\int x^2\mu(dx)-\int\int \log\left(|x-y|\right)\mu(dx)\mu(dy)-\frac{3}{4}$$

The next theorem is just inequality (7).

Theorem (3.2.5) [102]: Assume $Q:\mathrm{T}\to\mathbf{R}$ is a function so that $Q\left(e^{ix}\right)-\rho x^2$ is convex on $\mathbf{R}$ for a given $\rho>-1/4$. If μ_Q is a solution to the minimization problem

$$I^Q=\inf_{\mu\in\mathrm{P}(\mathrm{T})}\mathrm{E}^Q(\mu), \tag{15}$$

Where

$$\mathrm{E}^{Q}(\mathrm{V}) = \int Q(z)\mathrm{V}(dz) - \int\int_{\mathrm{T}\times\mathrm{T}} \log|z-z'|\mathrm{V}(dz)\mathrm{V}(dz'), \qquad (16)$$

then, for any $\mathrm{V} \in \mathrm{T}$, we have

$$(\rho + 1/4)\left(\mathrm{W}_2(\mathrm{V}, \mathrm{V}_Q)\right)^2 \le \mathrm{E}^{Q}(\mathrm{V}) - I^{Q}. \qquad (17)$$

In particular, there is a unique solution for the minimization problem (15).

Proof: We identify $[-\pi,\pi]$ with T via the exponential map $x \to e^{ix}$ and move the measure V to μ and V_Q to μ_Q. We then follow the proof of Theorem (3.2.2): with the necessary adjustments needed. The function $f(t)$ there becomes here

$$f(t) = -(\rho+1/4)t^2 \int |\theta(x)-x|^2 \mathrm{V}_Q(dx) + \int Q\left(e^{i(t\theta(x)+(1-t)x)}\right)\mathrm{V}_Q(dx)$$
$$-\int\int \log\left(\left|e^{i(t(\theta(x)+(1-t)x))} - e^{i(t\theta(y)+(1-t)(y))}\right|\right)\mathrm{V}_Q(dx)\mathrm{V}_Q(dy) - I^{Q}$$

Now, $\left|e^{ia}-e^{ib}\right|^2 = 4\sin^2((a-b)/2)$ for a,b real numbers and

$$\int |\theta(x)-x|^2 \mathrm{V}_Q(dx) = \frac{1}{2}\int\int \left((\theta(x)-x)-(\theta(y)-y)\right)^2 \mathrm{V}_Q(dx)\mathrm{V}_Q(dy)$$

Next, set $\theta_t(x) = t\theta(x) + (1-t)x$ and notice that.

$$\mathrm{g}(t) = -\frac{t^2}{4}\int |\theta(x)-x|^2 \mathrm{V}_Q(dx) - \int\int \log\left(\left|e^{it\theta_t(x))} - e^{i\theta_t(y))}\right|\right)\mathrm{V}_Q(dx)\mathrm{V}_Q(dy)$$
$$= -\int\int \frac{t^2}{8}\left((\theta(x)-x)-(\theta(y)-y)\right)^2 \mathrm{V}_Q(dx)\mathrm{V}_Q(dy)$$
$$-\int\int \log\left|2\sin\left((\theta_t(x)-\theta_t(y))/2\right)\right| v_Q(dx)\mathrm{V}_Q(dy)$$
$$= -2\int\int_{x>y} \frac{t^2}{8}\left((\theta(x)-x)-(\theta(y)-y)\right)^2 \mathrm{V}_Q(dx)\mathrm{V}_Q(dy)$$
$$-2\int\int_{x>y} \log\left(2\sin\left((\theta_t(x)-\theta_t(y))/2\right)\right)\mathrm{V}_Q(dx)\mathrm{V}_Q(dy)$$

where in the last line we used the fact that θ is a non-decreasing function. Since $x, y, \theta(x), \theta(y) \in [-\pi,\pi]$ and for $0 < a < b < \pi$, we have

$$\frac{d^2}{dt^2}\left(\frac{t^2}{8}(a-b)^2-\log\left(\sin\left(\frac{ta+(1-t)b}{2}\right)\right)\right) = \frac{(a-b)^2}{4}\left(\frac{1}{\sin^2\left(\frac{ta+(1-t)b}{2}\right)}-1\right)\geq 0$$

which implies that the function g is convex on [0,1]. This coupled with the convexity of $Q(e^{ix})-\rho x^2$ concludes that f is a convex function. Finally

$$f(t)\geq-(\rho+1/4)t^2\int\left|\theta(x)-x\right|^2 \mathrm{V}_Q(dx),$$

and thus, Lemma (3.2.1) : shows that $f(1)\geq 0$, which is (17).

The existence of a minimizer follows from the fact that E^Q is lower semicontinuous.

For $Q=0$ and $\rho=0$, the minimizer of (14) is the Haar measure on T . One can check this by showing directly that the uniform measure satisfy the variational form of (15).

Corollary (3.2.6) [102]: For any $\mu\in p(\mathrm{T})$

$$\frac{1}{4}\left(\mathrm{W}_2\left(\mu,\frac{dx}{2\pi}\right)\right)^2\leq-\iint_{\mathrm{T}\times\mathrm{T}}\log\left|z-z'\right|\mu(dz)\mu(dz').$$

Using the same argument as in the proof of Theorem (3.2.2): we can also prove a discrete version of it.

Theorem (3.2.7) [102]: Let $Q:\mathbf{R}\to\mathbf{R}$ be a function so that $Q(x)-\rho x^2$ is convex for a certain $\rho>0$. For $X=(x_1,x_2,...,x_n)\in\mathbf{R}^n$, set the energy of X to be given by

$$\mathrm{E}_n^Q(X)=\frac{1}{n}\sum_{k=1}^{n}Q(x_i)-\frac{2}{n(n-1)}\sum_{1\leq i\leq j\leq n}\log\left|x_i-x_j\right|$$

If $\Delta_n^Q=\mathrm{E}_n^Q(y)=\inf\left\{\mathrm{E}_n^Q(X):x\in\mathbf{R}^n\right\}$, then for any $X\in\mathbf{R}^n$,

$$\rho\left(\mathrm{W}_2(\mu(X),\mu(y))\right)^2\leq E_n^Q(X)-E_n^Q(y)=E_n^Q(X)-\Delta_n^Q \tag{18}$$

where $\mu(X)=\frac{1}{n}\sum_{k=1}^{n}\delta_{xk}$ Moreover,

$$\Delta_n^Q\leq\Delta_{n+1}^Q \tag{19}$$

The only statement that needs to be clarified here is (19). If y_{n+1} is a minimum point for E_{n+1}^{Q} and y_{n+1}^{i} denotes the n dimensional vector obtained from y_{n+1} by removing the i th component, then $\Delta_{n+1}^{Q} = \frac{1}{n+1}\sum_{i=1}^{n+1} E_n^Q\left(y_{n+1}^i\right)$ which is obviously $\geq \Delta_n^Q$.

The minimum points of E_n^Q are called Fekete points in the literature. It is known (see [206]) that $\lim_{n\to\infty}\Delta_n^Q = I^Q$, with I^Q defined in (10). We will reprove this fact below in Proposition (3.2.8) For $Q(x)=x^2$, the formula [168, 11] with the appropriate scaling gives the formula for computing $\Delta_n = \Delta_n^Q$ as

$$\Delta_n = \frac{1}{2}\left(1+\log(n-1)\right) - \frac{1}{n(n-1)}\sum_{j=1}^{n} j\log j = \frac{1}{2} - \frac{\log n}{n-1} - \frac{1}{n}\sum_{j=1}^{n-1}\frac{j}{n-1}\log\left(\frac{j}{n-1}\right) \qquad (20)$$

The next statement is a similar result to Theorems (3.2.2): and (3.2.7):

Proposition (3.2.8) [102]: Assume $Q:\mathbf{R}\to\mathbf{R}$ is a function so that $Q(x)-\rho x^2$ is convex for a certain $\rho>0$. Then for any $\mathrm{V}\in p(\mathbf{R})$ and $y\in\mathbf{R}^n$ a Fekete point for E_n^Q, we have

$$\rho\left(\mathrm{W}_2\left(\mathrm{V},\mu(y)\right)\right)^2 \leq \mathrm{E}^Q(\mathrm{V}) - \Delta_n^Q \qquad (21)$$

Furthermore, if μ_Q is the minimizing measure of E^Q, and $y_n\in\mathbf{R}^n$ is a Fekete point for E_n^Q, then

$$\lim_{n\to\infty}\Delta_n^Q = I^Q \text{ and } \lim_{n\to\infty}\mathrm{W}_2\left(\mu_Q,\mu(y_n)\right)=0, \qquad (22)$$

hence, $\mu(y_n)\xrightarrow[n\to\infty]{}\mu_Q$ weakly.

Proof: In the first place there is nothing to prove if $\mathrm{E}^Q(\mathrm{V})=\infty$. Therefore we assume that $\mathrm{E}^Q(\mathrm{V})<\infty$. Integrating (17) with respect to $\mathrm{V}(dx_1)\mathrm{V}(dx_2)\ldots\mathrm{V}(dx_n)$, one gets that

$$\rho\int\left(\mathrm{W}_2\left(\mu(X),\mu(y)\right)\right)^2\mathrm{V}(dx_1)\mathrm{V}(dx_2)\ldots\mathrm{V}(dx_n)\leq \mathrm{E}^Q(\mathrm{V})-\Delta_n^Q.$$

We finish the proof of (21) by showing the

$$\int\left(\mathrm{W}_2\left(\mu(X),\mu(y)\right)\right)^2\mathrm{V}(dx_1)\mathrm{V}(dx_2)\ldots\mathrm{V}(dx_n) = \left(\mathrm{W}_2\left(\mathrm{V},\mu(y)\right)\right)^2. \qquad (23)$$

To do this, we proceed by induction. For $n=1$, this statement becomes

$$\int \left(\mathrm{W}_2\left(\delta_x,\delta_y\right)\right)^2 \mathrm{V}(dx) = \left(\mathrm{W}_2\left(\mathrm{V},\delta_y\right)\right)^2$$

which, (9), is equivalent to the following (here F_{V} is the cumulative distribution function of V) .

$$\int |x-y|^2 \mathrm{V}(dx) = \int_0^1 \left|y - F_{\mathrm{V}}^{-1}(t)\right|^2 dt \, .$$

This can be checked by changing the variable in the second integral.

Assume (23) is true for $n-1, n \geq 2$. A simple application of (9) gives that

$$\left(\mathrm{W}_2\left(\mu(X),\mu(y)\right)\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left|x_{\sigma(i)} - y_{\tau(i)}\right|^2,$$

where σ and τ are permutations

of $\{1,2,...,n\}$ so $x_{\sigma(1)} \leq x_{\sigma(2)} ... \leq x$ If we denote by X_i the vector X with the ith component removed and similarly for y_i, one deduces

$$\left(\mathrm{W}_2\left(\mu(X),\mu(y)\right)\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{W}_2\left(\mu(X_i),\mu(y_i)\right)\right)^2 \tag{24}$$

On the other hand,

$$\left(\mathrm{W}_2\left(\mathrm{V},\mu(y)\right)\right)^2 = \sum_{k=0}^{n-1}\int_{k/n}^{(k+1)/n}\left|y_{\tau(k)} - F_{\mathrm{V}}^{-1}(t)\right|^2 dt$$

which can be used to argue that

$$\left(\mathrm{W}_2\left(\mathrm{V},\mu(y)\right)\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{W}_2\left(\mathrm{V},\mu(y_i)\right)\right)^2 \tag{25}$$

Putting together (24) and (25) and the induction hypothesis one finishes the proof of (23).To prove (22), we first point out that (21) applied to μ_Q yields that $I^Q \geq \Delta_n^Q$ for any $n \geq 1$. In particular this means that Δ_n^Q is bounded. Since $-\log|x-y| \geq -\frac{\rho}{4}\left(x^2+y^2\right)+c$ for a certain constant c, we get that $\Delta_n^Q \geq \frac{\rho}{4n}\sum_{i=1}^{n} x_i^2 - C$, where C is a constant. This implies that the sequence $\left\{\int x^2 \mu(y_n)(dx)\right\}_{n\geq 1}$ is bounded, whose consequence is that the sequence of measures $\mu(y_n)$ is tight, therefore there is a weak convergent subsequence $\mu(y_{nk})$ to a measure V. Now, for any $L > 0$, we have

$$\int \min\left\{\left(\left(Q(x)+Q(y)\right)/2-\log|x-y|\right),L\right\}\mu\left(y_{nk}\right)(dx)\mu\left(y_{nk}\right)(dy)\le \Delta_n^Q+L/nk$$

and this demonstrates that for any $L>0$,

$$\int \min\left\{\left(\left(Q(x)+Q(y)\right)/2-\log|x-y|\right),L\right\}v(dx)v(dy)\le I^Q$$

and, after passing $L\to\infty$, this yields

$$\mathrm{E}^Q(\mathrm{V})\le I^Q .$$

This together with (19) and the uniqueness of μ_Q from Theorem (3.2.2): ends the proof of $\lim_{n\to\infty}\Delta_n^Q=I^Q$. The rest follows.

In this section we deal with β -ensembles, which are studied in- [98]. These are tridiagonal matrices with independent entries of the form

$$A_n=\frac{1}{\sqrt{\beta n}}\begin{bmatrix} N(0,2) & \chi(n-1)\beta & & & \\ \chi(n-1)\beta & N(0,2) & \chi(n-2)\beta & & \\ & \ddots & \ddots & \ddots & \\ & & \chi 2\beta & N(0,2) & \chi\beta \\ & & & \chi\beta & N(0,2) \end{bmatrix}.$$

Here $N(0,2)$ stands for a normal with mean 0 and variance 2, while χ_γ is the

χ -distribution with parameter γ. The joint distribution of the eigenvalues is

$$\frac{1}{Z_{\beta,n}}\prod_{1\le i<j\le n}\left|x_i-x_j\right|^\beta \exp\left(-\beta n\sum_{i=1}^n x_i^2\right)$$

where here $Z_{\beta,n}$ is a normalization constant.

Set $\mu_n=\sum_{k=1}^n \delta_{\lambda k,n}$ the empirical distribution of the eigenvalues $\left\{\lambda_{k,n}\right\}_{k=1}^n$ of A_n

Theorem (3.2.8) [102]: Set $E_n=\frac{1}{2n}\sum_{k=1}^n\lambda_k^2-\frac{2}{(n-1)n}\sum_{1\le j<k\le n}\log|\lambda_i-\lambda_i|$ the energy of the eigenvalues $\{\lambda_k\}_{k=1}^n$ of A_n. If Δ_n is the quantity defined in (21), then almost surely,

$$\lim_{n\to\infty} n\left(\mathrm{E}_n-\Delta_n\right)=\Psi(1+\beta/2)-\log(\beta/2), \qquad (26)$$

where $\psi(x) = \frac{d}{dx}\log\Gamma(x)$ and $\Gamma$is the Gamma function. In addition, we have that

$$n^{1/2}\left(n\left(E_n - \Delta_n\right) - \left(\psi\left(1+\beta/2\right) - \log\left(\beta/2\right)\right)\right) \xrightarrow[n\to\infty]{} N\left(0, \psi'\left(1+\beta/2\right)\right), \quad (27)$$

where the convergence is in distribution sense.

The large deviations of $n\left(E_n - \Delta_n\right)$ is governed by the rate function

$$R^*(t) = \sup\left\{tz - R(z) : z \in \mathbf{R}\right\},$$

$$R(z) = \begin{cases} z + \left(\beta/2 - z\right)\log\left(\beta/2 - z\right) - \log\left(\dfrac{\Gamma\left(1+\beta/2-z\right)}{\Gamma\left(1+\beta/2\right)}\right) - \left(\beta/2\right)\log\left(\beta/2\right), & z < \beta/2 \\ \infty & z \geq \beta/2. \end{cases}$$

Proof: The proof is based on a version of Selberg's formula and elementary approximations involving Gamma function.

First, we have

$$\mathrm{E}\left[\exp\left(z\mathrm{E}_n\right)\right] = \frac{\int_{\mathbf{R}^n}\prod_{1\le i<j\le n}\left|x_i - x_j\right|^{\beta - \frac{2z}{n(n-1)}}\exp\left(-\beta n - \frac{z}{2n}\right)\sum_{j=1}^n x_j^2\,dx}{\int_{\mathbf{R}^n}\prod_{1\le i<j\le n}\left|x_i - x_j\right|^{\beta}\exp\left(-\beta n\sum_{j=1}^n x_j^2\right)dx}$$

and then, as a consequence of Selberg's formula [168], we get for complex z, that

$$\mathrm{E}\left[e^{z\mathrm{E}_n}\right] = \begin{cases} \dfrac{\left(n\beta/2 - z/n\right)^{-\frac{n}{2}\left[(n-1)\left(\beta/2 - \frac{z}{n(n-1)}\right)+1\right]}\prod_{j=1}^n \dfrac{\Gamma\left(1 + j\left(\beta/2 - \frac{z}{n(n-1)}\right)\right)}{\Gamma\left(1+\beta/2\right)}}{\left(n\beta/2\right)^{-\frac{n}{2}\left[(n-1)\beta/2+1\right]}\prod_{j=1}^n \dfrac{\Gamma\left(1+j\beta/2\right)}{\Gamma\left(1+\beta/2\right)}}, & \Re(z) < \beta/2 \\ \infty, & \Re(z) < \beta/2. \end{cases}$$

We need Stirling formula for approximation of Gamma function in the following form

$$\log\Gamma(t+1) = \left(t + 1/2\right)\log t - t + \log\left(2\pi\right)/2 + \mathrm{O}\left(\frac{1}{1+t}\right) \quad \text{for } t \geq 0$$

Using this and the above formula for $\mathrm{E}\left[\exp\left(z\mathrm{E}_n\right)\right]$ and (20), after some arrangements we get

$$\log\left(\mathrm{E}\left[e^{z(E_n-\Delta_n)}\right]\right)=\frac{z}{n-1}+\frac{z}{2}\log\left(1+\frac{1}{n-1}\right)$$

$$-\frac{z}{n-1}\log\left(\frac{\beta}{2}-\frac{z}{n^2}\right)+\frac{z(n+1)}{2(n-1)}\log\left(1+\frac{z}{n\left[(n-1)n\beta/2-z\right]}\right)$$

$$+\frac{n\left[(n-1)\beta+1\right]}{2}\log\left(1-\frac{z}{(n-1)\left[n^2\beta/2-z\right]}\right)+\frac{n\beta}{2}\log\left(1-\frac{2z}{n(n-1)\beta}\right)$$

$$-n\left[\log\left(1+\frac{\beta}{2}-\frac{z}{n(n-1)}\right)-\log\left(1+\frac{\beta}{2}\right)\right]+\mathrm{O}\left(\frac{z}{n^2}\right)$$

From this, replacing z by nz, one immediately obtains that for any $z\in\mathbf{R}$,

$$\log\left(\mathrm{E}\left[\exp\left(zn\left(E_n-\Delta_n\right)\right)\right]\right)\xrightarrow[n\to\infty]{}z\frac{\Gamma'(1+\beta/2)}{\Gamma(1+\beta/2)}-z\log(\beta/2)=z\left(\psi(1+\beta/2)-\log(\beta/2)\right)$$

Applying(27) with z replaced by $n^{\frac{3}{2}},z$, one can prove that for any complex z

$$\log\left(\mathrm{E}\left[\exp\left(zn^{1/2}\left(n\left(E_n-\Delta_n\right)-\left(\psi'(1+\beta/2)-\log(\beta/2)\right)\right)\right)\right]\right)\xrightarrow[n\to\infty]{}z^2\psi'(1+\beta/2)/2$$

whose consequence is (26). This, applied for $z=\pm1$ together with Chebyshev inequality yields

$$P\left(\left|n\left(\mathrm{E}_n-\Delta_n\right)-\left(\psi'(1+\beta/2)-\log(\beta/2)\right)\right|\geq\epsilon\right)\leq Ce^{-\varepsilon n^{1/2}}$$

for a certain constant $C>0$. This and an application of Borel-Cantelli's Lemma prove (26). Again applying (25)with n^2z in place of z, we can show that

$$\frac{1}{n}\log\left(\mathrm{E}\left[\exp\left(zn^2\left(\mathrm{E}_n-\Delta_n\right)\right)\right]\right)\xrightarrow[n\to\infty]{}R(z)$$

for any $z\in\mathbf{R}$. As a consequence of standard large deviations results (see in [111]) we conclude the proof of the last part of the theorem.

Chapter 4

Ricci Curvature for Metric Measure Spaces

The definitions are in terms of the displacement convexity of certain functions on the associated Wasserstein metric space of probability measures. We show that these properties are preserved under measured Gromov-Hausdorff limits. We give geometric and analytic consequences. Moreover, in the converse direction discretizations of metric measure spaces with curvature greater than or equals to the real constant will have rough curvature greater than or equal to the real number. We apply our results to concrete examples of homogenous planar graphs. We show a length of successive maps in a closed unit interval. We generalize the perturbations related to the Wasserstein distance.

Section(4.1): Geometry and Functionals of Wasserstein space

In this section we first recall some facts about convex function ,we then define gradient norms length space and measured Gromov-Hausdorff convergence. Finally, we define the 2- Wasserstein metric W_2 on $\mathrm{P}(X)$.

Let us recall a few results from convex analysis. (See [231])

Given a convex lower semi continuous function $U:\mathbf{R}\to\mathbf{R}\cup\{\infty\}$ (which we assume is not identically ∞), its Legendre transform $U^*:\mathbf{R}\to\mathbf{R}U\{\infty\}$ is defined by

$$U^*(p)=\sup_{r\in R}\left[pr-u(r)\right] \qquad (1)$$

Then U^* is also convex and lower semi continuous. We will sometimes identity a convex lower semi continuous. Function U define on a closed interval $I\subset\mathbf{R}$ with the convex function defined on the whole of $\mathbf{R}$ by extending U by ∞ outside of I .

Let $U[0,\infty)\to\mathbf{R}$ be a convex lower semi continuous function. Then U admits a left derivative $U_-:(0,\infty)\to\mathbf{R}$ and a right derivative $U:\ [0,\infty]\to\{-\infty\}\cup\ \mathbf{R}$, with $U(0,\infty)\subset\mathbf{R}$.

Furthermore, $U_-^{'}\le U_+'$ .They agree almost everywhere and are both non-decreasing. We will write

$$U'(\infty)=\lim_{r\to\infty}U_+'(r)=\lim_{r\to\infty}\frac{U(r)}{r}\in\mathrm{R}\cup\{\infty\} \qquad (2)$$

If we extend U by ∞ on $(-\infty,0)$ then its Legendre transform $U:\mathbf{R}\to\mathbf{R}\cup\{\infty\}$ becomes $U(p)=\sup_{r>0}(r)]$It is non-decreasing in P , infinite on $(U(\infty),\infty)$ and equals $-U(0)$ on $(-\infty,U(0)]$.Furthermore it is Continuous on $[-\infty,U(\infty))$.For all $r\in[0,\infty)$,we have $U^*(U(r))=rU'_+(r)-U(r)$.Let (X,d)be a compact metric space (with d valued in $[0,\infty)$).

Then Open ball of radius r around $\mathrm{x}\in\mathrm{X}$will be denoted by $\mathrm{B}_r(x)$and the sphere of radius r around x will be denoted by $\mathrm{S}_r(x)$.

Let $L^\infty(X)$denote the set of bounded measurable function on X.(We will consider such a function to be defined every where).Let $Lip(X)$ denote the set of Lipschitz functions on X . Given $f\in\mathrm{lip}(x)$, we define the gradient norm of f by

$$|\nabla f|(x)=\lim_{y\to x} sup\frac{[f(y)-(x)]}{d(x,y)} \tag{3}$$

If x is not an isolated point, and$|\nabla f|(x)=0$if x is isolated then $|\nabla f|\in\mathrm{L}^\infty(X)$. On some occasions will use a finer notion of gradient norm:

$$|\nabla^- f|(x)=\lim_{y\to x}\sup\frac{[f(y)-f(x)]_-}{d(x,y)}:=\lim_{y\to x}\sup\frac{[f(y)-(x)]_+}{d(x,y)} \tag{4}$$

If X is not isolated, and $|\nabla f|(x)=0$ if X is isolate. Here $\mathrm{a}_+=\max(\mathrm{a},0)$and $\mathrm{a}_-=\max(-\mathrm{a},0)$.

Clearly$|\nabla^- f|(x)\le|\nabla f|(x)$.note that $|\nabla^- f|(x)$is automatically zero if f has a local minimum at$X$. In a sense, $|\nabla^- f|(x)$ measures the downward pointing component of F near x .

If γ is curve in X, i.e a continuous map$\gamma:[0:1]\to x$, this its length is

$$L(\gamma)=\sup_{J\in N}\ \sup_{0=to\le t_1\le\ldots\le t_J=1}\sum_{J=1}^{J} d\left(\gamma\left(t_{j-1}\right),\gamma\left(t_j\right)\right). \tag{5}$$

Clearly $L(\gamma)\ge d(\gamma(0),\gamma(1))$.

We will assume that X is a length space, meaning that the distance between two points $x_0,x_1\in X$ is the infirmum of the length of curve from $x_0$to x_1 .Such a spade is path connected.

As X is compact, it is a strictly in transit length space ,meaning that we can replace infirmum by minimum [29] That Is for any $x_0, x_1 \in X$ there is minimal geodesic (possible non-unique) from x_0 to x_1.We may sometimes write geodesic instead of "minimal geodesic".

By [29], any minimal geodesic γ joining x_0 to x_1 can be parameterized uniquely by $t \in [0,1]$ so that

$$d\left(\gamma(t),\gamma(t')\right) = |t-t'|\, d\left(x_0, x_1\right) \quad (6)$$

We will often assume that the geodesic has been so parameterized .By definition a subset $A \subset X$ is convex if for any $x_0, x_1 \in A$ there is a minimizing geodesic from x_0 to x_1 that lies entirely in A. It is totally convex if for any $x_0, x_1 \in A$,any minimizing geodesic in x from x_0 to x_1 lies in A. Given $\lambda \in \mathbf{R}$ a function $F : X \to \mathbf{R}$ and only $t \in [0,2]$ we have

$$F(\gamma(t)) \le yF(\gamma(1)) + (1-t)F(\gamma(0)) - \frac{1}{2}\lambda t(1-t)L(\gamma)^2 \quad (7)$$

In the case when x is a smooth Riemannian manifold with Riemannian metric g, and $F \in C^2(x)$, this is the same as saying that $hessF \ge \lambda g$.

Definition (4.1.1) [121]: Given two compact metric spaces (x_1, d_2) and (x_2, d_2)) an

ε-Gromov- Hausdorff approximation from x_1 to x_2 is a (not necessarily continuous) map $f : X_1 \to X_2$ so that

(i) $\left| d_2 F(x_i), F(x_1^1) - d_1(x_1, x_1') \right| \le \varepsilon, \forall x_1, x_1 \in X_1,$

(ii)For all $x_2, \in X_2$, there is an $x_1 \in X_1$ so that $d_2\left(F(x_2), x_2\right) \le \varepsilon$.

An $\varepsilon -$ Gromov-Hausdorff approximation $f : X_2 \to X_2$ has an approximate inverse $F : X_2 \to X_1$ which can be constructed as follows: Given $x_2 \in X_2$ choose $x_1 \in X_1$ so that $d_2(f(x_1), x_3) \ge \varepsilon$ and put $f'(x_2) = x_1$ then f' is a $3\varepsilon -$ Gronov-Hausdorff approximation from X_2 to X_1.Moreover, for all $x_1 \in X_1, d_1(\alpha_1, (f' \circ f)(x)) \in 2\varepsilon$,and for all $x_2 \in X_2, d_2(x_2(f \circ f')(x_2) \le \varepsilon)$,

Definition (4.1.2) [121]: A Sequence of compact metric spaces $\{X_i\}_{i=1}^{\infty}$ converges to X in the Gromov-Hausdorff topology if there is a sequence of ε_i –approximations $f_i : X_i \to X$ with $\lim_{i\to\infty} \varepsilon_i = 0$

This notion of convergence comes from a metrizable topology on the space f all compact metric spaces modulo isometries. If $\{X_i\}_{i=1}^{\infty}$ are length spaces that converge to X in Gromov-Hausdorff topology. Then X is also a length space [29]. For the purpose of this section , we can and will assume that maps f and f' in Gronov-Hausdorff approximation are Borel probably measures on X. We give $P(X)$ the weak−*topology, i.e

$\lim_{i\to\infty} \mu_i = \mu$, if and only if for all $F \in C(X)$, $\lim_{i\to\infty} \int_{X_i} Fd\mu = \int_X Fd\mu$. (8)

Definition (4.1.3) [121] Given $V \in P(X)$.Consider this metric-measure space (X, d, V). A sequence $\{X_i, d_i, V_i\}_{i=1}^{\infty}$ converge to (X, d, V) in the measured Gronov-Hausdorff topology, if there are ε_i - approximations are $f_i : X_i \to X$ with $\lim_{i\to\infty} \varepsilon_i = 0$ so that $\lim_{i\to\infty} (f_i)^* V_i = V \; in \; P(X)$.

Other topologies on the class of metric-measure spaces are discussed in [166]. For later use we note the following generalization of the Arzela-Ascoli Theorem.

Lemma (4.1.4) [121] : [163] Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of compact metric spaces converging to X in the Gromov-Hausdorff topology with ε_i-approximations $f_i : X_i \to X$ .Let $\{Y_i\}_{i=1}^{\infty}$ be a sequence of compact metric spaces converging to Y in the

Gromov-Hausdorff topology with ε_i -approximations $g_i : Y_i \to Y$. For each i, let $f_i' : X \to X_i$ be an approximate inverse to f_i as in the paragraph following Definition (4..1.1): Let $\{\alpha\}_{i=1}^{\infty}$ be a sequence of maps $\alpha_i : X_i \to Y_i$ that are asymptotically equicontinuous in the sense that for every $\varepsilon > 0$, there are $\delta = \delta(\varepsilon) > 0$ and $N = N(\varepsilon) > \square^+$ that for all $i \geq N$

$$dX_i\left(x_i, x_i^1\right) < \delta \Rightarrow dy_i\left(\alpha_i\left(x_i\right), \alpha_i\left(x_i^1\right)\right) < \varepsilon \qquad (9)$$

Then after passing to a subsequence to maps $g_i \circ \alpha_i \circ f_i' X \to y$ converge uniformly to a continuous map $\alpha : X \to Y$.

In the conclusion of Lemma (4.1.4): the maps $g_i \circ \alpha_i \circ f_i'$ may not be continuous, but the notion of uniform convergence makes sense nevertheless .

Given $\mu_0, \mu_1 \in \mathrm{P}(X)$ we say that a probably measure $D \in \mathrm{P}(X \times X)$, is a transference plan between μ_0 and μ_1 if.

$$(p_0)_* \Pi = \mu_0, (p_1)_* \Pi = \mu_1 \qquad (10)$$

Where $p_0, p_1 : X \times X \to X$ are projection onto the first and second factors, respectively. In wards Π represents a way to transport the mass from μ_0 to μ_1 and $\Pi(x_0, x_2)$ is the a moment of mass which is taken from appoint x_0. And brought to a point x_2

We will use optimal transport with quadrate cost function (sequence of the distance). Namely, given $\mu_0, \mu_1 \in \mathrm{P}(x)$, we consider the variational problem.

$$\mathrm{W}_2(\mu_0, \mu_1)^2 = \inf_{\lambda} \int_{X \times X} d(x_0, x_1)^2 d\Pi(x_0, x_1) \qquad (11)$$

Where π ranges over the set of all transference plans between μ_0 and μ_1. Any minimizer π for this variational problem is called an optimal transference plan.

In (11), one can replace the infimum by the minimum [23], i.e. there always exist (at least) one optimal transference plan. Since X has finite diameter, the infimum is obviously finite . The quantity W_2 will be called the Wasserstein distance of order 2 between μ_0 and μ_1 , it defines a metric on $\mathrm{P}(X)$, the topology that it induces on $\mathrm{P}(X)$ is the weak -*topology [23]. When equipped with the metric W_2 , $\mathrm{P}(X)$, is a compact metric space, which we will often denote by $\mathrm{P}_2(X)$, we remark that there is an isometric embedding $X \to \mathrm{P}_2(X)$,given by $x \to \delta x$. This shows that diam $(\mathrm{P}_2(X)) \ge diam(X)$. Since the reverse inequality follows from the definition of W_2 actually. $(\mathrm{P}_2(X)) = \mathrm{diam}(X)$, a monge transport plan coming from a map $F : X \to X$ with $F_* \mu_0 = \mu_1$ given by $\pi = (Id,, F)_* \mu_0$,. In general an optimal transference plan does not have to be a monge transport . Although this may be true under some assumption

A function $\phi : X \to [-\infty, \infty]$, is $\frac{d^2}{2}$ concave if it is not identically $-\infty$ and it can be written in the form

$$\phi(x) = \inf_{x' \in X} \left(\frac{d(x,x')^2}{2} - \tilde{\phi}(x') \right) \tag{12}$$

For some function $\tilde{\phi} : X \to [-\infty, \infty)$, such functions play an important role in the description of optimal transport on Riemannian manifolds.

In this section. We investigate some features of the Wasserstein space $P_2(X)$ associated to a compact length space (X,d). (Recall that the subscript 2 in $P_2(X)$ means that $P(X)$ is equipped with the 2-Wasserstein metric). We show that $P_2(X)$ is a length space. We define displacement interpolation and show that every Wasserstein geodesic comes from a displacement interpolation. We then recall some fact about optimal transport on Riemannian manifolds.

We denote by $\mathrm{Lip}([0,1], X)$, the space of Lipschitz continuous maps $C:[0,1] \to X$ with the uniform topology. For any $k > 0$

$$\mathrm{Lip}_k([0,1], X) = \{C \in Lip([0,1], X) : d(C(t)C(t')) \le k|t-t'| \; for\ all\ t, t' \in [0,1]\} \tag{13}$$

is a compact subset of $Lip([0,1], X)$.

Let Γ denote the set of minimizing geodesies on X. It is closed subspace

$$\mathrm{Lip}_{diam(X)}([0,1], X),$$

defined by equation $L(C) : d(C(0), C(1))$. For any $t \in [0,1]$, the evolution map $e_t : \Gamma \to X$ defend by

$$e_t(\gamma) = \gamma(t) \tag{14}$$

is continuous. Let $E : \Gamma \to X \times X$ be the “endpoint” map given by $E(\gamma) = (e_1(\gamma))$, A dynamical transference plan consists of a transference plan π and a Borel measure Π on Γ such that $E_* \Pi = \pi$; it is said to be optimal if Π itself is. In words the transference plan π tells us how mach mass goes from a pint x_0 to another point x_1, but does not tell us about the actual path that the mass has to follow. Intuitively, mass should follow a long geodesies, but there may be several possible choices of geodesies between two given points and the transport may be divided among these geodesies

,this is the information provided by Π ,but there may be several π . If Π is on optimal dynamical transference plan, then for $t \in [0,1]$ we out

$$\mu_t = (\ell_t)_* \Pi \tag{15}$$

The one-parameter family of measures $\{\mu_t\}_{t\in[0,1]}$ is called displacement interpolation in wards μ_t is what has becomes of the mass of μ_0 after it has traveled from time 0 to time t according to the dynamical transference plan Π .

Lemma (4.1.5) [121]: The map $c c[0,1] \to \mathrm{P}_2(X)$ given by $c(t) = \mu_t$ has length $L(c) = \mathrm{W}_2(\mu_0, \mu_1)$.

Proof: Given $O \le t \le t' \le \mathrm{I}, (e_t, e_{t'})_* \Pi$ is a particular transference plan from μ_t to $\mu_{t'}$ and so

$$\mathrm{W}_2(\mu_t, \mu_{t'}) \le \int_{X\times X} d(x_0, x_1)^2 d\left((e_t, e_{t'})_* \Pi\right)(x_0, x_1)$$

$$= \int_\Gamma d(\gamma(t), \gamma(t'))^2, d\Pi(\gamma) = \int_\Gamma (t'-t)^2 L(\gamma)^2 d\Pi(\gamma)$$

$$= (t'-t) \int_{X\times X} d(x_0, x_1)^2 dE_* \Pi(x_0, x_1) = (t'-t)^2 \mathrm{W}_2(\mu_0, \mu_1)^2 \tag{16}$$

Equation (15) implies that $L(c) \le \mathrm{W}_2(\mu_0, \mu_1)$ and so $L(c) \equiv \mathrm{W}_2(\mu_0, \mu_1)$

Proposition (4.1.6) [121]: Let (X, d) be a compact length space then any two point $\mu_0, \mu_1 \in p_2(X)$ can be joined by a displacement interpolation .

corollary (4.1.7)[121] , If X is a compact length space then $\mathrm{P}_2(X)$ is a compact length space. ..

Example (4.1.8) [121] : Suppose that $X = A \cup B \cup C$ where A, B and C are subsets of the plane given by $A = \{(x_1, 0) : -2 \le x_1 \le -1\}$, $B = \{(x_1, x_2) : x_1^2 + x_2^2 = 1\}$ and $C = \{(x_1, 0) : I \le x_1 \le 2\}$. Let μ be the one –dimentional Hausdorff measure of A and let μ_1 be the one - dimentional Hausdorff measure of C . Then there is an uncountable number of Wasserstein geodesies from μ_0 to μ_1 given by the whims of a switchman at the point $(-1, 0)$.

Corollary(4.1.9)[274]:The map $c c[0,1] \to \mathrm{P}_2(X)$ given by $c(t) = \mu_t$ has length $L(c) = \mathrm{W}_2(\mu_0, \mu_1)$.

Proof: Given $0 \le t, t+\varepsilon \le \mathrm{I}, (e_t, e_{t+\varepsilon})_*$ and Π is a transference plan from μ_t to $\mu_{t+\varepsilon}$ and so

$$\mathrm{W}_2(\mu_t, \mu_{t'}) \le \int_{X\times X} d(x_0, x_1)^2 d\left((e_t, e_{t+\varepsilon})_* \Pi\right)(x_0, x_1)$$

$$= \int_\Gamma d(\gamma(t), \gamma(t+\varepsilon))^2, d\Pi(\gamma) = \int_\Gamma \varepsilon^2 L(\gamma)^2 d\Pi(\gamma)$$

$$= \varepsilon \int_{X\times X} d(x_0, x_1)^2 dE_* \Pi(x_0, x_1) = \varepsilon^2 \mathrm{W}_2(\mu_0, \mu_1)^2. \qquad (17)$$

Hence equation (15) gives the result.

The next result states that every Wasserstein geodesies arises from a displacement interpolation .

Proposition (4.1.10) [121]: Let (X,d) be a compact length space and let $\{\mu_t\}_{t\in[0,1]}$ is the displacement interpolation associated to Π.

Proof: Let $\{\mu_t\}_{t\in[0,1]}$ be a Wasserstein geodesies U_p up to reparametrization , we can assume that for all $t, t' \in [0,1]$

$$\mathrm{W}_2(\mu_t, \mu_{t'}) = |t - t'| \mathrm{W}_2(\mu_0, \mu_1) \qquad (18)$$

Let $\pi x_0^{(o)} x_{1/2}$ be an optimal transference plan from μ_0 to $\mu_{1/2}$ and let $\pi_{x_{1/2}}^{1/2}, x_1$ be optimal transference plan from $\mu_{1/2}$ to μ_1 .Consider the measure obtained by gluing together $\pi x_0^{(o)} x_{1/2}$ and $\pi_{x_{1/2}}^{1/2}, x_1$.

$$M^{(1)} = \frac{d\pi_{x_0}^{(0)} x_{1/2} d\pi x_{1/2}^{(1/2)}, x_1}{d\mu_{1/2}(x_{1/2})} \qquad (19)$$

on $X \times X \times X$. The precise meaning of this expression is just as in the ''gluing Lemma'' started in [23]: Decompose $\pi^{(o)}$ with respect to the projection $p_1 : X \times X \to X$ on the second factor as $\Pi^{(o)} = \sigma x_{1/2}^{(o)} \mu_{1/2}(x_{1/2})$, where for $\mu_{1/2}$-almost all $x_{1/2}, \sigma x_{1/2}^{(o)} \in p\left(p_1^{=1}(x_{1/2})\right)$ is a probability measure on $p_0^{-1}(x_{1/2})$. Decompose $\Pi^{(1/2)}$ with respect to the projection $p_o : X \times X \to X$ on the first factor as $\Pi^{1/2} = \sigma x_{1/2}^{(1/2)} \mu_{1/2}(x_{1/2})$ where for $\mu_{1/2}$-almost all $x_{1/2}, \sigma x_{1/2}^{(1/2)} \in \mathrm{P}\left(p_o^{-1}(x_{1/2})\right)$. Then for $F \in c(X \times X \times X)$

$$\int_{X\times X\times X} F dM^{(1)}$$

$$= \int_X \int_{p_1^{-1}(x_{1/2}) \times p_0^{-1}(x_{1/2})} F(x_0, x_{1/2}, x_1) \, d\sigma x_{1/2}^{(o)} x d\sigma x_{1/2}^{(1/2)}(x_1) \, d\mu_{1/2}(x_{1/2}) \tag{20}$$

The formula

$$d\pi x_0, x_1 = \int_X M^{(1)}_{x_0, x_{1/2}, x_1} \tag{21}$$

Defines a transference plan from μ_0 to μ_1 with cost

$$\int_{X \times X} d(x_0, x_1)^2 d\pi x_0, x_1 \leq \int_{X \times X \times X} d(x_0, x_{1/2}) + d(x_{1/2}, x_1) \frac{d\pi^{(o)} x_0, x_{1/2}}{d\mu_{1/2}(x_{1/2})} d\pi^{1/2} x_{1/2}, x_1$$

$$\leq \int_{X \times X \times X} 2\left(d(x_0, x_{1/2})^2 + d(x_{1/2}, x_1)^2\right) \frac{d\pi^{(1/2)} x_0, x_{1/2} d\pi^{1/2} x_{1/2}, x_0}{d\mu_{1/2}(x_{1/2})}$$

$$\leq 2\left(\int_{X \times X} d(x_0, x_{1/2})^2 d\pi^{(0)}_{x_0, x_{1/2}} + \int_{X \times X} d(x_{1/2}, x_1)^2 d\pi^{(1/2)}_{x_{1/2}, x_1}\right)$$

$$= 2\left(\mathrm{W}_2(\mu_0, \mu_{1/2})^2 + \mathrm{W}_2(\mu_{/2}, \mu_1)^2\right) = \mathrm{W}_2(\mu_0, \mu_1)^2. \tag{22}$$

This π is an optimal transference plan and we must have equality every where in (21). Let

$$B^{(1)} = \left\{(x_0, x_{1/2}, x_1) \in X \times X \times X : d(x_0, x_{1/2}) = dx_{1/2}, x_1 = \frac{1}{2} d(x_0, x_1)\right\}; \tag{23}$$

Then $M^{(1)}$ is supported on $B^{(1)}$. For $t \in \{0, 1/2\}$, define $e_t : B^{(1)} \to X$ by $e_t(x_0, x_{1/2}, x_1) = x_t$. Then $(e_t)_* M^{(1)} = \mu_t$

We can repeat the same procedure using a decomposition of the interval [0,1] into 2^i subintervals. For any $i \geq I$ define .

$$B^{(i)} = \left(x_0, x_2^{-i}, x_{22}^{-1}, \ldots, x_{1-2}, x_1\right) \in X^{2i+1} \tag{24}$$

$$d(x_0, x_{2^{-i}}) = d, (x_{2^{-i}}, x_{2.2^{-i}}) = \ldots = d(x_{1-2^{-i}}, x_1) = 2^{-i} d(x_0, x_1)$$

For $0 \leq j \leq 2^i - 1$ choose an optimal transference plan $\pi^{(j, 2^{-i})}_{x_{j,2^{-i}}, x_{(j+1)2^{-i}}}$ from $\mu^{-i}_{j.2}$ to $\mu_{(j+1), 2^{-1}}$.Then as before we obtain a probability measure $M^{(i)}$ on $B^{(i)}$ by

$$M^{(i)}_{x_0, x_{2-i}}, \ldots . x_1 = \frac{d\pi^{(0)}(x_0, x_{2^{-i}}) d\pi^{(2^{-i})}(x_{2^{-i}}, x_{2.2^{-i}}) \ldots d\pi^{(1-2^{-i})}(x_{1-2^{-i}}, x_1)}{d\mu_{2^{-i}}(x_{2^{-i}}) \ldots d\mu_{1-2^{-i}}(x_{1-2^{-i}})} \tag{25}$$

The formula

$$d\pi_{x_0,x_1} = \int_{X^{2^i_{-1}}} M^{(i)}_{x_0,x_{2^{-i}},\dots,x_1} \tag{26}$$

Defines a transference plan from μ_0 to μ_1. For $t = j.2^{-i}, 0 \le j \le 2^i$ define $e_t : B^{(1)} \to X$ by $e_t\left(x_0,\dots,x_{1/2},x_1\right) = x_t$;then $\left(e_t\right)_* M^{(c)} = \mu_i$.

Let S be in the proof of Proposition (4.1.6): Given $\left(x_0,\dots,x_1\right) \in B^{(i)}$ define a map $p_{x_0,x_1,\dots x_1} : [0,1] \to X$, as the concatenation of the paths $S\left(x_0,x_{2^{-1}}\right), S\left(x_{2^{-i}},x_{2.2^{-i}}\right),\dots,$ and

$S\left(x_{1-2^{-i}},x_1\right)$. As $p_{x_0,\dots,x_1}$ is normalized continuous curve from x_0. to x_1 length $d\left(x_0,x_1\right)$ it is a geodesies. For each the linear function L on $C(\Gamma)$ given by

$$F \to \int_{x^{2^i+1}} F\left(p_{x_0,\dots,x_1}\right) dM^{(i)}_{x_0,\dots,x_1} \tag{27}$$

Define a probability measure $\mathbf{R}^{(i)}$ on the compact space Γ. Let $\mathbf{R}^{\infty}$ be the limit of a weak -* convergent subsequence of $\left\{\mathbf{R}^{(i)}\right\}_{i=1}^{\infty}$ it is also a probability measure on Γ.

For any $t \in \frac{N}{2^v} \cap [0,1]$ and $f \in c(X)$ we have $\int_k \left(e_t\right)_* f dR^{(i)} = \int_X f d\mu_t$ for large I. Then $\int_k \left(e_t\right)_* f dR^{(\infty)} = \int_X f d\mu_t$ for all $f \in c(X)$, or equivalently ,$\left(e_t\right)_* \mathbf{R}^{(\infty)} = \mu t$. But as in the proof of Lemma (4.1.5) $\left(e_t\right)_* R^{(\infty)}$ is weak-* continuous in t. It follows that $\left(e_t\right)_* \mathbf{R}^{(\infty)} = \mu_t$ for all $t \in [0,1]$

We discuss the case when X is a smooth compact connected Riemannian manifold M with Riemannian metric g.(The results are also valid if. G is only C^3 smooth). Given $\mu_0,\mu_1 \in \mathrm{P}_2(M)$ which are absolutely continuous with respect to $dvol_M$ it is known that there is a unique Wasserstein geodesies c joining μ_0 to μ_1 [242]. Furthermore; for each $t \in [0,t], c(t)$ is absolutely continuous with respect to $dvol_M$ [31]. Thus it makes sense to talk about the length space $\mathrm{P}_2^{ac}(M)$ of Borel probability measures on M that are absolutely continuous with respect to the Riemannian density equipped with the metric W_2. It is a dense totally convex subset of $\mathrm{P}_2(M)$. Note that if M is other than a point a dense totally, then $\mathrm{P}_2^{ac}(M)$ is an incomplete metric space and it is neither open nor closed in $\mathrm{P}_2(X)$.An optimal transfer-ence plan in $\mathrm{P}_2^{ac}(M)$ turn out to be monge transport that is $c(t) = \left(F_i\right)_* \mu_0$ for a family of

Monge transport $\{F_t\}_{t\in[0,1]}$ of M. For each $m \in M$, $F_t(m) = \exp_m(-t\nabla\phi(m))$

[31]. This function ϕ just as any $\frac{d^2}{2}$ – concave function on a compact Riemannian manifold, is Lipschitz [242] and has Hessian every where [31]. If we only want the Wasserstein geodesies to be defined for an interval $\left[0,r^{-1}\right]$ then we can use the same formula with ϕ being $r\frac{d^2}{2}$-concave.

All of our results will involve a distinguished reference measure, which is not a prioric canonically given. So by ''measured length space '' we will mean a triple (X,d,V), where $(X.d)$ is a compact length space and V is a Borel probability measure on X. These assumptions automatically imply that V is a regular measure we write.

$$\mathrm{P}_2(X,\mathrm{V}) = \{\mu \in \mathrm{P}_2(X) : \mathrm{supp}(\mu) \subset \mathrm{supp}(\mathrm{V})\} \tag{28}$$

We note by $\mathrm{P}_2^{ac}(X,\mathrm{V})$ the elements of $\mathrm{P}_2(X,\mathrm{V})$ that are absolutely continuous with respect to V .

Definition (4.1.11) [121]: Let $\cup$ be a continuous convex function on $[0,\infty]$ with $\cup(0)=0$. Given $\mu,\mathrm{V}\in\mathrm{P}_2(X)$, we define the functional $U_{\mathrm{V}} : \mathrm{P}_2(X) \to \mathbf{R}\cup\{\infty\}$ by

$$U_{\mathrm{V}}(\mu) = \int_X U(\rho(X))\, d\mathrm{V}(x) + U'(\infty)\mu_s(X), \tag{29}$$

Where

$$\mu = \rho\mathrm{V} + \mu_s \tag{30}$$

is the Llebesgue decomposition of μ with respect to V into an absolutely continuous part $\rho\mathrm{V}$ and a singular part μ_s ,we have the

$$\int_X U(\rho(x))\, d\mathrm{V}(x) \ge U\left(\int_X \rho(x)\, d\mathrm{V}(x)\right). \tag{31}$$

Lemma (4.1.12) [121]: $U_{\mathrm{V}}(\mu) \ge U_{\mathrm{V}}(\mathrm{V}) = U(1)$.

Proof: as $\cup$ is convex for any $\alpha \in (0.1)$ we have

$$U(\alpha r + 1 - \alpha) \le \alpha U(r) + (1-\alpha)U(1) \tag{32}$$

$$U(r)-U(1)\geq\frac{1}{\alpha}\left[U(\alpha r+1-\alpha)-U(1)\right] \tag{33}$$

Then

$$\int_X U(\rho)\,d\mathrm{V}-U(1)\geq\int_X\frac{U(\alpha\rho+1-\alpha)-U(1)}{\alpha\rho-\alpha}(p-1)\,d\mathrm{V} \tag{34}$$

Where we take the integrand f the right-hand-side to vanish at points $x\in X$ where $\rho(x)=1$. We break up the right-side of (34) according to whether $\rho(x)\leq 1$ or $\rho(x)>1$ from monotone convergence for $\rho\geq 1$ we have

$$\lim_{\alpha\to 0^+}\int_X\frac{U(\alpha\rho+1-\alpha)-U(1)}{\alpha\rho-\alpha}(\rho-1)I_{\rho\leq 1}d\mathrm{V}=U'_-(1)\int_X(\rho-1)1_{\rho\leq 1}d\mathrm{V}$$

While for $\rho>1$ we have

$$\lim_{\alpha\to 0^+}\int\frac{U(\alpha\rho+1-\alpha)-U(1)}{\alpha\rho-\alpha}(\rho-1)I_{\rho>1}d\mathrm{V}=U'_+(1)\int_X(\rho-1)1_{\rho>1}d\mathrm{V} \tag{35}$$

$$\int_X U(\rho)d\mathrm{V}-U(1)\geq U'_-(1)\int_X(\rho-1)d\mathrm{V}+\left(U'_+(1)-U'_-(1)\right)\int_X(\rho-1)I_{\rho>1}d\mathrm{V} \tag{36}$$

As $U_v(\mathrm{V})=U(1)$ the Lemma follows. □

Definition (4.1.13) [121]: Given a compact measured length space (X,d,V) and a number $\lambda\in R$, we say that U_V is.

(i) λ-displacement convex if for all Wasserstein geodesies $\{\mu_t\}_{t\in[0.1]}$ with $\mu_0,\mu_1\in P_2(X,\mathrm{V})$, we have

$$U_v(\mu_t)\leq tU_v(\mu_1)+(1-t)U_v(\mu_0)-\frac{1}{2}\lambda t(1-t)W_2(\mu_0,\mu_1)^2 \tag{37}$$

for all $t\in[0,1]$

(ii)weakly λ-displacement convex if for all $\mu_0,\mu_1\in P_2(X,\mathrm{V})$, there is some Wasserstein geodesies from μ_0 to μ_1 along which (37) is satisfied

weakly $\lambda-a.c$. displacement convex if the condition is satisfied when we just assume that $\mu_0,\mu_1\in P_2^{ac}(X,\mathrm{V})$.

If U_V is λ-displacement convex and $\operatorname{supp} U = X$, then the nction $t \to U_V(\mu_t)$ is λ-convex on [0,1],i.e. for all $0 \le s \le s' \le 1$ and $t \in [0,1]$.

$$U_V\left(\mu_{ts'+(1-t)s}\right) \le tU_V(\mu_{s'}) + (1-t)U_V(\mu_s) - \frac{1}{2}\lambda t(1-t)(s'-s)W_2(\mu_0,\mu_1)^2 \qquad (38)$$

This is not a priori the case if we only assume that U_V is weakly ƛ displacement convex .

$$\begin{array}{ccc} \lambda\text{-displacement convex} & \Rightarrow & \text{weakly } \lambda\text{-displacement convex} \\ \Downarrow & & \Downarrow \\ \lambda\text{-a.c displacement convex} & \Rightarrow & \text{weakly } \lambda\text{-a.c displacement convex} \end{array} \qquad (39)$$

The next proposition reverse the right vertical implication in (39)

Proposition (4.1.15) [121]: Let U be a continuous convex function on $[0,\infty]$ with $U(0)=0$. Let (X,d,V) be compact measured length space. Then U_V is weakly

λ-displacement convex, if and only if it is weakly $\lambda - a.c$.displacement convex.

Proof : We must show that if U_V is weakly

$\lambda - a.c$ displacement convex, then it is weakly λ-displacement convex, that is for $\mu_0, \mu_1 \in P_2(X,V)$, we must show that there is some Wasserstein geodesies $\{\mu_t\}_{t\in[0,1]}$ from μ_0 to μ_1 along which

$$U_V(\mu_t) \le tU_V(\mu_0) + (1-t)U_V(\mu_1) - \frac{1}{2}\lambda t(1-t)W_2(\mu_0,\mu_1)^2 \qquad (40)$$

We may assume that $U_V(\mu_0) < \infty$ and $U_V(\mu_1) < \infty$ as otherwise (38) is trivially true for any Wasserstein geodesies from μ_0 to μ_1 . There are sequences $\{\mu_{k,0}\}_{k=1}^{\infty}$ and $\{\mu_{k,1}\}_{k=1}^{\infty}$ in $P_2^{ac}(X,V)$ (in fact with continuous densities so that $\lim_{k\to\infty} \mu_{k,0} = \mu_0, \lim_{k\to\infty} \mu_{k,1} = \mu_1$, $\lim_{k\to\infty} U_V(\mu_{k,0}) = U_V(\mu_0)$ and $\lim_{k\to\infty} U_V(\mu_{k,1}) = U_V(\mu_1)$

Let $c_k = [0,1] \to P_2(X)$ be a minimal geodesies from $\mu_{k,0}$, to $\mu_{k,1}$ such that for all $t \in [0,1]$.

$$U_V(C_k(t)) \le tU_V(\mu_{k,1}) + (1-t)U_V(\mu_{k,0}) - \frac{1}{2}\lambda t(1-t)W_2(\mu_{k,0},\mu_{k,1})^2 \qquad (41)$$

After taking a subsequence, we may assume that the geodesies $\{C_k\}_{k=1}^{\infty}$ converge uniformly $\left(i.e.in C\left([0,1], \mathrm{P}_2(X)\right)\right)$ to geodesies $c:[0,1]\rightarrow \mathrm{P}_2(X)$ from μ_0 to μ_1 [29]. The lower semi-continuity of U_{V}, implies that $U_{\mathrm{V}}(c(t)) \le \liminf_{k\to\infty} U_{\mathrm{V}}(c_k(t))$.

The proposition follows. In fact the proof of Proposition (4.1.15): gives the following slightly stronger result.

Lemma (4.1.16) [121]: Let U be a continuous convex function on $(0,\infty)$ with $U_0 = 0$. Let (X, d, V) be a compact measured length space. Suppose that for all $\mu_0, \mu_1 \in \mathrm{P}_2^{ac}(X, \mathrm{V})$, with continuous densities, there is some Wasserstein geodesies from μ_0 to μ_1 along with (22) is so satisfied. Then U_{V} is weakly $\lambda -$displacement convex. The next lemma gives sufficient conditions for the horizontal implications in (39) to be reversed . We recall the definition of total convexity.

Lemma (4.1.17) [121]: (i) Suppose that X has the property that for each minimizing geodesies $C = [0,1] \rightarrow \mathrm{P}_2(X)$, where is some $\delta_c > 0$ so that minimizing geodesies between $C(t)$ and $C(t')$ is unique whenever $|y - t'| < \delta c$. Suppose that $\mathrm{supp}(\mathrm{V}) = X$. If U_{V} is weakly $\lambda -$displacement convex.

(ii) Suppose that $\mathrm{P}_2^{as}(X', \mathrm{V})$, is totally convex in $\mathrm{P}_2(X)$. Suppose that X has properly that for each minimizing geodesies $C:[0,1] \rightarrow \mathrm{P}_2^{ac}(X, \mathrm{V})$, there is some $\delta_c > 0$ so that the minimizing geodesies between $c(t)$ and $c(t')$ is unique whenever $|t - t'| \le \delta_c$, Suppose that $\mathrm{supp}\,(\mathrm{V}) = X$. If U_{V} is weakly λ-displacement convex, then it is $\lambda - a.c$ displacement convex,

The following functional will play an important role.

Definition (4.1.18) [121]: Put

$$U_N(\gamma) = \begin{cases} Nr\left(1 - r^{-1/N}\right) & if\ I < N < \infty \\ r \log r & if\ N = \infty \end{cases} \tag{42}$$

Definition (4.1.19) [121]: Let $H_{N,\mathrm{V}} : p_2(X) \rightarrow [0,\infty]$ be the functional associated to U_{V} via definition (4.1.11)]: More explicitly.

-For $N \in (1,\infty)$

$$H_{N,V} = N - N\int_X p^{1-1/N} dV \tag{43}$$

Where ρV is the absolutely continuous part in the Lebesgue decomposition of μ with respect to V.

-For $N=\infty$ the functional $H_{\infty,V}$ is defined as follows : If μ is absolutely continuous with respect to V with $\mu=\rho V$ then

$$H_{\infty,V}(\mu) = \int \rho \log \rho dV \tag{44}$$

While if μ is not absolutely continuous with respect to V then $H_{\infty,V}(\mu)=\infty$. To verify that $H_{N,V}$ is indeed the functional associated to U_N we note that $U'_N(\infty)=N$

And write.

$$N\int_X \rho\left(1-\rho^{-1/N}\right) dV + N\mu_s(X) = N\int_X \rho\left(1-\rho^{-1/N}\right) dV + N\left(1-\int_X \rho dV\right)$$

$$= N - N\int_X \rho^{1-1/N} dV. \quad . \tag{45}$$

Of course the deference of treatment of the singular part of V according to whether N is finite or not reflects the fact that U_N grows at most linearly when $N<\infty$ but super linearly when $N=\infty$,ensures that $H_{N,V}$ is lower semicontinuous on $P_2(X)$.

Definition (4.1.20) [121]: Let (X,d,V) be a compact measured length space. Let U be a continuous convex function on $[0,\infty]$ with $U(0)=o$ which is C_-^2 regular on $[0,\infty]$. Given $\mu\in P_2^{ac}(X,V)$ with $\rho=d\mu/dV$ a positive Lipsehitz function on X, define the generalized Fisher information I_U by .

$$I_U(\mu) = \int_X U^n(\rho)^2 \left|\nabla^-\rho\right|^2 d\mu = \int_X \rho U''(\rho)^2 \left|\nabla^-\rho\right|^2 dV. \tag{46}$$

The following estimated generalize the ones that underline the HWI inequalities in [69].

Proposition (4.1.21) [121]: Let (X,d,V) be a compact measured length space. Convex function Let U be a continuous convex function on $[0,\infty]$

with $U(0)=0$.Given $\mu \in P_2(X,V)$. Let $\{\mu_t\}_{t\in[0.1]}$ be a Wasserstein geodesies from $\mu_0=\mu$ to $\mu_1=V$. Given $\lambda \in R$, suppose that (35) is satisfied. Then

$$\frac{\lambda}{2} W_2(\mu,V)^2 \le U_V(\mu)-U_V(V). \qquad (47)$$

Now suppose in addition that U is C_-^2 regular on $(0,\infty)$ and that $\mu \in P_2^{ac}(X,V)$ is such that $\rho=\frac{d\mu}{dV}$ is a positive Lipsechitz function on X. Suppose that $U_V(\mu)<\infty$ and $\mu_t \in P_2^{ac}(X,U_V)$, for all $t\in[0,1]$.Then

$$U_V(\mu)-U_V(V) \le W_2(\mu,V)\sqrt{I_U(\mu)}-\frac{\lambda}{2} W_2(\mu,V)^2 \qquad (48)$$

Proof: consider the function $\phi(t)=U_V(\mu_t)$. Then $\phi(0)=U_V(\mu)$ and $\phi(1)=U_V(V)$. By assumption,

$$\phi(t) \le t\phi(1)+(1-t)\phi(0)-\frac{1}{2}\lambda t(1-t) W_2(\mu,V)^2 \qquad (49)$$

If $\phi(0)-\phi(1)<\frac{1}{2}\lambda W_2(\mu,V)^2$ then $\phi(t)-\phi(1)\le(1-t)\left(\phi(0)-\phi(1)-\frac{1}{2}\lambda W_2(\mu,V)^2\right)$, we conclude that $\phi(t)-\phi(1)$ is negative for t close to 1, which contradicts Lemma (4.1.12): This $\phi(0)-\phi(1)\ge\frac{1}{2}\lambda W_2(\mu,V)$, which proves (46).

To prove (47) put $\rho_t=\frac{d\mu t}{dv}$, then $\phi(t)=\int_X U(\rho_t)dV$. From (48)for $t>0$ we have

$$\phi(0)-\phi(1) \le -\frac{\phi(t)-\phi(0)}{t}-\frac{1}{2}\lambda(1-t) W_2(\mu,V)^2 \qquad (50)$$

To prove the inequality (48), it suffices to prove that

$$\liminf_{t\to 0}\left(-\frac{\phi(t)-\phi(0)}{t}\right) \le W_2(\mu,V)\sqrt{I_U(\mu)} \qquad (51)$$

The convexity of U implies that

$$U(\rho_t)-U(\rho_0) \ge U'(\rho_0)(\rho_t-\rho_0) \qquad (52)$$

Integrating with respect to V and dividing by $t<0$, we infer

$$-\frac{1}{t}\left[\phi(t)-\phi(0)\right] \le -\frac{1}{t}\int_X U'\left(\rho_0(x)\right)\left[d\mu_t(x)-d\mu_0(x)\right] \tag{53}$$

By Proposition (4.1.10): $\mu_t = (e_t)_* \Pi$ where Π is a certain probability measure on the space Γ of minimal geodesies in X In particular,

$$-\frac{1}{t}\int_X U'\rho_0(x)\left[d\mu_t(x)-d\mu_0(x)\right] = -\frac{1}{t}\int_\Gamma \left[U'\left(\rho_0(\gamma(t))\right)-U'\left(\rho_0(\gamma(0))\right)\right] d\Pi(\gamma) \tag{54}$$

Since U' is non-decreasing and $td\left(\gamma(0),\gamma(1)\right)=d\left(\gamma(0),\gamma(t)\right)$, we have

$$-\frac{1}{t}\int_\Gamma [U'\left(\rho_0(\gamma(t))\right)-U'\left(\rho_0(\gamma(0))\right)]d\Pi(\gamma) \le -\frac{1}{t}\int_\Gamma I_{\rho_0(\gamma(t))\le\rho_0(\gamma(0))} U'\left(\rho_0(\gamma(t))\right)-U'\left(\rho_0(\gamma(0))\right)d\Pi(\gamma),$$

$$= \int_\Gamma \frac{U'\left(\rho_0(\gamma(t))\right)-U'\left(\rho_0(\gamma(0))\right)}{\rho_0(\gamma(t))-\rho_0(\gamma(0))} \frac{\left[\rho_0(\gamma(t))-\rho_0(\gamma(t))\right]_-}{d\left(\gamma(0),\gamma(t)\right)} - d\left(\gamma(0),\gamma(0)\right)d\pi(\gamma) \tag{55}$$

Where strictly speaking we define the integrand of the last term to be zero when

$$\rho_0\left(\gamma(t)\right) = \rho_0\left(\gamma(0)\right)$$

Applying the Cauchy-Schwarz inequality, we can bound the last term above

$$\sqrt{\int_\Gamma \frac{\left[U'\left(\rho_0(\gamma(t))\right)-U'\left(\rho_0(\gamma(0))\right)\right]^2}{\left[\rho_0(\gamma(t))-\rho_0(\gamma(0))\right]^2} \frac{\left[\rho_0(\gamma(t))-\rho_0(\gamma(0))\right]_-^2}{d\left(\gamma(0),\gamma(t)\right)^2} d\Pi(\gamma)} \sqrt{\int_\Gamma d\left(\gamma(0),\gamma(I)\right)^2 d\Pi(\gamma)} \tag{56}$$

The second square root is just $W_2(\mu_0,\mu_1)$. To conclude the argument, it suffices to show that

$$\liminf_{t\to 0} \int_\Gamma \frac{\left[U'\left(\rho_0(\gamma(t))\right)-U'\left(\rho_0(\gamma(0))\right)\right]^2}{\left[\rho_0(\gamma(t))-\rho_0(\gamma(0))\right]^2} \frac{\left[\rho_0(\gamma(t))-\rho_0(\gamma(0))\right]_-^2}{d\left(\gamma(0),\gamma(0)\right)^2} d\Pi(\gamma) \le I_U(\mu) \tag{57}$$

The continuity of ρ_0 implies that $\lim_{t\to 0} \rho_0\left(\gamma(t)\right) = \rho_0\left(\gamma(0)\right)$.So

$$\lim_{t\to 0}\int_\Gamma \frac{\left[U'\left(\rho_0(\gamma(t))\right)-U'\left(\rho_0(\gamma(0))\right)\right]^2}{\left[\rho_0(\gamma(t))-\rho_0(\gamma(0))\right]^2} = U''\left(\rho_0(\gamma(0))\right)^2 \tag{58}$$

On the other hand , the definition of the gradient implies

$$\limsup_{t\to 0} \int_\Gamma \frac{\left[\rho_0(\gamma(t))-\rho_0(\gamma(0))\right]_-^2}{d\left(\gamma(0),(\gamma(t))\right)^2} \le \left|\nabla^- \rho\right|^2\left(\gamma(0)\right). \tag{59}$$

As ρ is a positive Lipschitz function on X and U' is also C^1-regular on $(0,\infty)$, $U_0'\rho_0$ is also Lipschitz on X. Then $\frac{\left[U'(\rho_0(\gamma(t)))-U'(\rho_0(\gamma(0)))\right]^2}{d(\gamma(0))-(\gamma(t))^2}$ is uniformly bounded on Γ, with respect to t, and dominated convergence implies that

$$\liminf_{t\to 0}\int_\Gamma \frac{\left[U'(\rho_0(\gamma(t)))-U'(\rho_0(\gamma(0)))\right]^2}{\left[\rho_0(\gamma(t))-\rho_0(\gamma(0))\right]^2}\frac{\left[\rho_0(\gamma(t))-\rho_0(\gamma(0))\right]^2}{d(\gamma(0),\gamma(t))^2}-d\pi(\gamma)\le I_U(\mu)$$

$$\le\int_\Gamma U''(\rho_0(\gamma(0)))^2\left|\nabla^-\rho_0\right|^2(\gamma(0))d\Pi(\gamma)=\int_X U''(\rho_0(x))^2\left|\nabla^-\rho_0\right|^2(x)d\mu(x). \qquad (60)$$

This concludes the proof of the inequality on the right-hand side of (49).

Particular case (4.1.22) [121]: Taking $U=U_N$ with $\mu=\rho v$ and $\rho\in Lip(X)$ appositive function define

$$I_{N,v}(\mu)=\begin{cases}\left(\frac{N-1}{N}\right)^2\int_X\frac{\left|\nabla^-\rho\right|^2}{\rho^{\frac{2}{N}+1}}dv & if\ I<N<\infty\\ \int_X\frac{\left|\nabla^-\rho\right|^2}{\rho}dv & ip\ N=\infty\end{cases} \qquad (61)$$

Proposition (4.1.23) [121]: implies the following inequalities :

-If $\lambda>0$ then

$$\frac{\lambda}{2}\mathrm{W}_2(\mu,\mathrm{V})^2\le H_{N,\mathrm{V}}(\mu)\le \mathrm{W}_2(\mu,\mathrm{V})\sqrt{I_{N,\mathrm{V}}(\mu)}-\frac{\lambda}{2}\mathrm{W}_2(\mu,\mathrm{V})^2\le\frac{1}{2\lambda}I_{N,\mathrm{V}}(\mu) \qquad (62)$$

-If $\lambda\le 0$

$$H_{N,v}(\mu)\le diam(X)\sqrt{I_{N,I}(\mu)}-\frac{\lambda}{2}diam(X)^2 \qquad (63)$$

Corollary (4.1.24) [121]: If a sequence of compact metric spaces $\{(X_i,d_i)\}_{i=1}^\infty$ converges (X,d) then $\{\mathrm{P}_2(X_i)\}_{i=1}^\infty$ converges in the Gromov-Hausdorff topology to $\mathrm{P}_2(X)$

Proposition (4.1.25) [121]: If $f:(X_1,d_1)\to(X_2,d_2)$ is an $\varepsilon-$Gromov-Hausdorff approximation then $f_*:\mathrm{P}_2\to\mathrm{P}_2(X_2)$ is an $\tilde{\varepsilon}$- Gromov-Hausdorff approximation where

$$\tilde{\varepsilon} = 4\varepsilon + \sqrt{3\varepsilon\left(2diam\left(X_2\right)+3\varepsilon\right)} \tag{64}$$

Proof: Given $\mu_1, \mu_1' \in \mathrm{P}_2\left(X_1\right)$, let π_1 be an optimal transference plan for μ_1 and μ_1' but $\pi_2 = \left(f \times f\right)_* \pi_1$. Then π_2 is a transference plan for $f_*\mu_1$ and $f_*\mu_1'$ we have

$$\mathrm{W}_2\left(f_*\mu_1, f_*\mu_1'\right) \leq \int_{X_2 \times X_2} d_2\left(x_2, y_2\right)^2 d\pi_2\left(x_2, y_2\right)$$

$$= \int_{X_1 \times X_1} d_2\left(f\left(x_1\right), f\left(y_1\right)\right)^2 d\Pi_1\left(x_1, y_1\right) \tag{65}$$

As

$$\left|d_2\left(f\left(x_1\right), f\left(y_1\right)\right)^2 - d_1\left(x_1, y_1\right)^2\right|$$

$$= \left|d_2 f\left(x_1\right), f\left(y_1\right) - d\left(x_1, y_1\right)\right| . \left(d_2 f\left(x_1\right), f\left(y_1\right) + d_1\left(x_1, y_1\right)\right) \tag{66}$$

We have

$$\left|d_2\left(f\left(x_1\right), f\left(y_1\right)\right)^2 - d_1\left(x_1, y_1\right)^2\right| \leq \varepsilon\left(2diam\left(X_1\right)+\varepsilon\right) \tag{67}$$

If follows that

$$\mathrm{W}_2\left(f_*\mu_1, f_*\mu_1'\right)^2 \leq \mathrm{W}_2\left(\mu_1, \mu_1'\right)^2 + \varepsilon\left(2diam\left(X_1\right)+\varepsilon\right) \tag{68}$$

and

$$\mathrm{W}_2\left(f_*\mu_1, f_*\mu_2'\right)^2 \leq \mathrm{W}_2\left(\mu_1, \mu_1'\right)^2 + \varepsilon\left(2diam\left(X_2\right)+\varepsilon\right) \tag{69}$$

If follows from this last inequality that

$$\mathrm{W}_2\left(f_*\mu_1, f_*\mu_1'\right)^2 \leq \mathrm{W}_2\left(\mu_1, \mu_1'\right) + \sqrt{\varepsilon\left(diam\left(X_2\right)\right)+\varepsilon} \tag{70}$$

We now exchange the roles of X_1 and X_2. We corresponding apply (68) instead of (69) to the map f' and the measures $f_*\mu_1$ and $f_*\mu_1'$ and use the fact f' is a

ε- Gromov-Hausdorff approximation, to obtain

$$\mathrm{W}_2 f_*'\left(f_*\mu_1\right), f_*'\left(f_*\mu_1'\right) \leq \mathrm{W}_2\left(f_*\mu_1, f_*\mu_1'\right) + \sqrt{3\varepsilon\left(2diam\left(X_2\right)+3\varepsilon\right)} \tag{71}$$

Since $f' \circ f$ is an admissible Monge transport between μ_1 and $(f' \circ f)_* \mu_1$ or between μ_1' and $(f' \circ f)_* \mu_1'$ which moves points by a distance at most 2ε we have

$$W_2\left((f' \circ f)_* \mu_1, \mu_1\right) \le 2\varepsilon, \quad W_2\left((f' \circ f)_* \mu_1', \mu_1'\right) \le 2\varepsilon \tag{72}$$

This by (70) and the triangle inequality,

$$W_2(\mu_1, \mu_1') \le W_2(f_*\mu_1, f_*\mu_1') + 4\varepsilon + \sqrt{3\varepsilon\left(2diam(X_2) + 3\varepsilon\right)} \tag{73}$$

Equation (70) and (73) show that condition (i) of Definition (4.1.1):is satisfied .

Finally, given $\mu_2 \in P_2(X_2)$ consider the Monge transport $f \circ f'$ from μ_2 to $(f \circ f')_* \mu_2$. Then $W_2\left(\mu_2, f_*(f_*'\mu_2)\right) \le \varepsilon$. Thus condition (ii) of Definition (4.1.1): is satisfied as well. □

Theorem (4.1.26) [121]: Let $\{X_i, d_i, V_i\}_{i=1}^{\infty}$ be a sequence of compact measure spaces so that $\lim_{i\to\infty}(X_i, d_i, V_i) = (X, d, V_\infty)$ in the measure Gromov-Hausderff topology . Let U be a continuous convex function on $[0,\infty)$ with $U(0)=0$. Given $\lambda \in R$, suppose that for all i,U_{V_i} is weakly λ-displacement convex for $\{X_i, d_i, V_i\}$. Then U_{V_∞} is weakly

λ-displacement convex for (X, d, V).

Proof : By Lemma (4.1.16): it surfaces to show that for any $\mu_0, \mu_1 \in P_2(X)$ with continuous densities with respect to V_∞ there is a Wasserstein geodesies joining them along which inequality (37) holds for V_∞ we may assume that $U_{V_\infty}(\mu_0) < \infty$ and $U_{V_\infty}(\mu_1) < \infty$ as otherwise any Wasserstein geodesies works.

Write $\mu_0 = \rho_0 V_\infty$ and $\mu_1 = \rho_1 V_\infty$. Let $f_i : X_i \to X$ be an ε-approximation, with $\lim_{\varepsilon\to\infty} \varepsilon_i = 0$ and $\lim_{i\to\infty}(fc)_* V_i = V_\infty$, if I is sufficiently large then $\int_X \rho_0 d(f_i)_* V_i > 0$ and $f_X \rho_1 d(f_i)_* V_i > 0$ for such i, put $\mu_{i,0} = \dfrac{(f_i^* \rho_0) V_i}{\int_X \rho_0 d(f_i)_* V_i}$ and $\mu_{i,1} = \dfrac{(f_i^* \rho_1) V_i}{\int_X \rho_1 d(f_i)_* V_i}$.

Then

$$\left(f^i\right)_* \mu_{i,0} = \frac{\rho_0 \left(f_i\right)_* \mathrm{V}_i}{\int_X \rho_0 d\left(f_i\right)_* \mathrm{V}_i} \text{ and } \left(f^i\right)_* \mu_{i,1} = \frac{\rho_1 \left(f_i\right)_* \mathrm{V}_i}{\int_X \rho_1 d\left(f_i\right)_* \mathrm{V}_i}.$$

Now choose geodesies $c_i : [0.1] \to \mathrm{P}_2\left(X_i\right)$ with $c_i(0) = \mu_{i,0}$ and $c_i(1) = \mu_{i,1}$ so that for all $t \in [0,1]$, we have

$$U_{\mathrm{V}_i}\left(C_i(t)\right) \leq tU_{\mathrm{V}_i}\left(\mu_i, 1\right) + (1-t)U_{\mathrm{V}_i}\left(\mu_i, 0\right) - \frac{1}{2}\lambda t(1-t)\mathrm{W}_2\left(\mu_{i.0}, \mu_{i,1}\right)^2. \tag{74}$$

From Lemma (4.1.4): and Corollary (4.1.24): after passing to a subsequence, the maps $(f_i)_* \circ c_i : [0,1] \to p_2(\mathrm{X})$ converge uniformly to a continuous map $c : [0,1] \to \mathrm{P}_2(X)$. As $\mathrm{W}_2\left(c_i(t), c_i(t')\right) = |t-t'| \mathrm{W}_2\left(\mu_{i,0}, \mu_{i,1}\right)$, it follows that $\mathrm{W}_2\left(c(t), c(t')\right) = |t-t'| \mathrm{W}_2\left(\mu_0, \mu_1\right)$. Thus C is a Wasserstein geodesic. The problem is to pass to the limit in (73) as $i \to \infty$.

Given $F \in c(x)$, the fact that $\rho_0 \in c(x)$ implies that

$$\lim_{i\to\infty} \int_X Fd\left(f_i\right)_* \mu_{i,0} = \lim_{i\to\infty} \int_X F\rho_0 \frac{d\left(f_i\right)_* \mathrm{V}_i}{\int_X \rho_0 d\left(f_i\right)_* \mathrm{V}_i} = \int_X F\rho_0 d\mathrm{V}_\infty. \tag{75}$$

Thus $\lim\left(f_i\right)_* \mu_{i,0} = \mu_0$ similarly , $\lim_{i\to\infty}\left(f_i\right)_* \mu_{i,1} = \mu_2$. It follows from Corollary (4.1.24): that

$$\lim_{i\to\infty} \mathrm{W}_2\left(\mu_{i,0}, \mu_{i,1}\right) = \mathrm{W}_2\left(\mu_0, \mu_1\right) \tag{76}$$

Next

$$U_{\mathrm{V}_i}\left(\mu_i, 0\right) = \int_{X_i} U\left(\frac{f_i^* \rho_0}{\rho_0 d\left(f_i\right)_* \mathrm{V}_i}\right) dv_i = \int_{X_i} U\left(\frac{\rho_0}{\int_X \rho_0 d\left(f_i\right)_* \mathrm{V}_i}\right) d\left(f_i\right)_* \mathrm{V}_i. \tag{77}$$

As

$$\lim_{i\to\infty} U\left(\frac{\rho_0}{\int_X \rho_0 d\left(f_i\right)_* \mathrm{V}_i}\right) = U\left(\rho_0\right) \tag{78}$$

Thus $\lim_{i\to\infty}\left(f_i\right)_* \mu_{i,0} = \mu_0$ .Similarly $\lim_{i\to\infty}\left(f_i\right)_* \mu_{i,2} = \mu_2$. It follows from Corollary (4.1.24): that

$$\lim_{i\to\infty} \mathrm{W}_2(\mu_{i,0}, \mathrm{V}_{i,1}) = \mathrm{W}_2(\mu_0, \mu_2) \tag{79}$$

Uniformly on X ,it follows that

$$\lim_{i\to\infty}\int_X U\left(\frac{\rho_0}{\int_X \rho_0 d(f_i)_* \mathrm{V}_i}\right)d(f_i)_* \mathrm{V}_i = \lim\int_X U(\rho_0)d(f_i)_* \mathrm{V}_i = \int_X U(\rho_0)d\mathrm{V}_\infty . \qquad (80)$$

Thus $\lim_{i\to\infty} U_{v_i}(\mu_{i,0}) = U_{v_\infty}(\mu_0)$.Similarly $\lim_{i\to\infty} U_{v_i}(\mu_{i,0}) = U_{v_\infty}(\mu_1)$. It follows that

$$U_{(f_i)_* v_i}\left((f_i)_* C_i(t)\right) \le U_{v_i}\left(c_i(t)\right). \qquad (81)$$

Then for any $t\in[0,1]$, we can combine this with the lower semicontinuity of $(\mu,\mathrm{V}) \to U_v(\mu)$ to obtain

$$U_{v_\infty}\left(c(t)\right) \le \liminf_{i\to\infty} U_{(f_i)_* v_i}\left((f_i)_* c_i(t)\right) \le \liminf_{i\to\infty} U_{v_i}\left(c_i(t)\right). \qquad (82)$$

Combining this with (76) and the preceding results, we can take $i\to\infty$ in (74) and find

$$U_{v_\infty}\left(c(t)\right) \le tU_{v_\infty}(\mu_1) + (1-t)U_{v_\infty}(\mu_0) - \frac{1}{2}\lambda t(1-t)W_2(\mu_0,\mu_1)^2$$

This concludes the proof.

Definition (4.1.27)[121]: Let F be family of continuous convex functions U on $[0,\infty)$ with $U(0)=0$. Given a function $\lambda : \mathrm{F} \to \mathbf{R}\cup\{-\infty\}$ we say a compact measured length space (X,d,V) is weakly λ-displacement convex for the family F if for any $\mu_0,\mu_1 \in \mathrm{P}_2(X,\mathrm{V})$, one can find a Wasserstein geodesic $\{\mu_t\}_{t\in[0,1]}$ is supposed to work for all of the functions $U\in\mathrm{F}$ .Hence if (X,d,V) is weakly λ-displacement convex for the family F then it is weakly λ(U)-displacement convex for each $U\in\mathrm{F}$, but the converse is not a priori true.

Theorem(4.1.28)[121]: Let $\{(X_i,d_i,\mathrm{V}_i)\}_{i=0}^{\infty}$ be sequence of compact measured length spaces with $\lim_{i\to\infty}(X_i,d_i,\mathrm{V}_i) = (X,d,\mathrm{V}_\infty)$ in the measured Gromov-Hausdorff topology. Let F be a family of continuous convex functions U on $[0,\infty)$ with $U(0)=0$. Given a function $\lambda : \mathrm{F} \to \mathbf{R}\cup\{-\infty\}$,suppose that each (X_i,d_i,V_i) is weakly λ-displacement convex for the family F .Then (X,d,V_∞) is weakly λ-displacement convex for the family F .

Proposition (4.1.29)[121]: Let be a family of continuous convex functions U on $[0,\infty)$ with $U(0)=0$. Given a function $\lambda : F \to \mathbf{R}\cup\{-\infty\}$, (X,d,V) is weakly λ-displacement convex for the family F if and only if it is weakly λ-a-c-displacement convex for the family F .

Section (4.2): Ricci Curvature for Measured Length Spaces And Riemannian Manifolds

This section deals with N-Ricci curvature and its basic properties. We first define certain classes DC_N of convex functions U. We use these to define the notions of a measured length space (X,d,V) having nonnegative N-Ricci curvature, or ∞-Ricci curvature bounded below by $K\in\mathbf{R}$ in [234].Consider a continuous convex function $U:[0,\infty)\to\mathbf{R}$ with $U(0)=0.$ We define the nonnegative function.

$$p(r)=rU_+'(r)-U(r), \tag{83}$$

with $p(0)=0$. If one thinks of U as defining an internal energy for a continuous $(0,\infty)$ then p can be thought of as a pressure. By analogy, if U is C^2-regular on $(0,\infty)$ then we define the "iterated pressure"

$$p_2(r)=rp'(r)-p(r). \tag{84}$$

Definition (4.2.1)[121]: For $N\in[1,\infty)$, we define DC_N to be the set of all continuous convex functions U on [0,1), with $U(0)=0$, such that the function

$$\phi(\lambda)=\lambda^N U\left(\lambda^{-N}\right) \tag{85}$$

is convex on $(0,\infty)$ .We further define DC_N to be the set of all continuous convex functions U on [0,1) , with $U(0)=0$ such that the function

$$\phi(\lambda)=e^{\lambda}U\left(e^{-\lambda}\right) \tag{85}$$

is convex on $(-\infty,\infty)$.

We note that the convexity of U implies that ϕ is non-increasing in λ, as $\frac{U(\alpha)}{\alpha}$ _is non-decreasing in x blow are some useful facts about the classes DC_N.

Lemma (4.2.2) [121]: If $N \le N'$ then $\mathrm{DC}_{N'} \subset \mathrm{DC}_N$.

Lemma (4.2.3) [121]: For $N \in [1,\infty]$,

(a) If U is a continuous convex function on $[0,\infty)$ with $U(0)=0$ then $U \in \mathrm{DC}_N$ if and only if the function $r \to p(r)/r^{1-\frac{1}{N}}$ if is non-decreasing on $(0,\infty)$

(b) If U is a continuous convex function on $[0,\infty)$ that is C^2-regular on $(0,\infty)$, with $U(0)=0$, then $U \in \mathrm{DC}_N$ if and only if $p_2 \ge -\frac{p}{N}$.

Proof]274]:(a) Suppose first that U is a continuous convex function on $[0,\infty)$ and $N \in [1,\infty)$.

Putting, $\psi(\lambda) = \lambda^N U(\lambda^{-N})$ $r = \lambda^{-N}$ *and* $\lambda = r^{\frac{-1}{N}}$ therefore $\psi(r^{\frac{-1}{N}}) = r^{-1}U(r)$.By Differentiating we get that

$$\psi_-^{'}(r^{\frac{-1}{N}}).(\frac{-1}{N}r^{\frac{-1}{N}-1}) = r^{-1}U^{'}(r) - \mathrm{r}^{-2}U(r) = \frac{1}{r}(rU^{'}(r) - U(r)).$$

So that

$$\psi_-^{'}(\lambda) = \frac{-N}{r^{\frac{-1}{N}-1}}.\frac{1}{r^2}(rU^{'}(r) - U(r)) = \frac{-N}{r^{1-\frac{1}{N}}}(rU^{'}(r) - U(r))$$

$$\psi_-^{'}(\lambda) = -Np(r)/r^{1-\frac{1}{N}}. \tag{87}$$

Then ψ is convex if and only if $\psi_-^{'}$ is non-decreasing, which is the case if and only if the function $r \mapsto p(r)/r^{1-\frac{1}{N}}$ is non-decreasing (since the map $\lambda \to \lambda^{-N}$ is non-increasing).

(b) Suppose that U is C_2-regular on $(0,\infty)$.We get

$$\psi^{''}(r^{\frac{-1}{N}}).(\frac{-1}{N}r^{\frac{-1}{N}-1}) = -\mathrm{N}\left(\frac{1}{r^{1-\frac{1}{N}}}(rU''(r) + U^{'}(r) - U^{'}(r)) - (1-\frac{1}{N})r^{-2+\frac{1}{N}}(rU^{'}(r) - U(r))\right)$$

and

$$\psi''(\lambda)=\frac{N^2}{r^{\frac{-1}{N}-1}}\left(\frac{1}{r^{\frac{-1}{N}}}\mathrm{U}''(\mathrm{r})-\frac{(1-\frac{1}{N})}{r^{\frac{2-\frac{1}{N}}{}}}(rU'(r)-U(r))\right)=\frac{N^2}{r^{\frac{-2}{N}+1}}\left(\mathrm{r}^2\mathrm{U}''(\mathrm{r})-p(r)+\frac{p(r)}{N}\right)$$

$$\psi''(\lambda)=N^2r^{\frac{2}{N}-1}\left(p_2(r)+\frac{p(r)}{N}\right). \tag{88}$$

Then ϕ is convex if and only if $\psi''\geq 0$, which is the case if and only if $p_2\geq -\frac{p}{N}$.

The proof in the case $N=\infty$ is similar.

Lemma (4.2.4)[121]: Given $U\in \mathrm{DC}_\infty$, either U is linear or there exist $a,b>0$ such that

$U(r)\geq a\,r\log r-br$

Proof: The function U can be reconstructed from ϕ by the formula

$$U(x)=x\phi(\log(1/x)). \tag{89}$$

As ϕ is convex and non-increasing, either ϕ is constant or there are constants $a,b>0$ such that $\phi(\lambda)\geq -a\lambda-b$ for all $\lambda\in\mathbf{R}$. In the first case, U is linear. In the second case, we have $U(x)\geq -ax\log(1/x)-bx$, as required.

We recall from Definition (4.1.27): the notion of a compact measured length space (X,d,V) being weakly λ-displacement convex for a family of convex functions F .

Definition (4.2.5)[121]: Given $N\in[1,\infty]$, we say that a compact measured length space (X,d,V) has nonnegative N -Ricci curvature if it is weakly displacement convex for the family DC_N.

By Lemma (4.2.2): if $N\leq N'$ and X has nonnegative N-Ricci curvature then it has non-negative N'- Ricci curvature. In the case $N=\infty$, we can define a more precise notion.

Definition (4.2.6)[121]: Given $K \in \mathbf{R}$, define $\lambda : DC_\infty \to \mathbf{R} \cup \{-\infty\}$ by

$$\lambda(U) = \inf_{r>0} K \frac{p(r)}{r} = \begin{cases} K \lim_{r \to 0^+} \frac{p(r)}{r} & if\ K > 0, \\ 0 & if\ K = 0, \\ K \lim_{r \to 0} \frac{p(r)}{r} & if\ K < 0, \end{cases} \qquad (90)$$

where p is given by (1). We say that a compact measured length space (X, d, V) has $\infty -$Ricci curvature bounded below by K if it is weakly A-displacement convex for the family DC_∞

If $K \le K'$ and (X, d, V) has $\infty -$-Ricci curvature bounded below by K' then it has

$\infty -$Ricci curvature bounded below by K.

The next proposition shows that our definitions localize on totally convex subsets.

Proposition (4.2.7)[121]: Suppose that a closed set $A \subset X$ is totally convex. Given $V \in P_2(X)$ with $V(A) > 0,$, put $V' = \frac{1}{V(A)} V|_A \in P_2(A)$..

(a) If (X, d, V) has nonnegative N-Ricci curvature then (A, d, V') has nonnegative

N -Ricci curvature.

(b) If (X, d, V) has ∞ -Ricci curvature bounded below by K then (A, d, V') has ∞ -Ricci curvature bounded below by K.

Proof: By Proposition (4.1.10) $P_2(A)$ is a totally convex subset of $P_2(X)$. Given $\mu \in P_2(A) \subset P_2(X)$, let $\mu = \rho V + \mu_s$ be its Lebesgue decomposition with respect to V. Then $\mu = \rho' V' + \mu_s$ is the Lebesgue decomposition of μ with respect to V', where $\rho' = V'(A) \rho|_A$. Given a continuous convex function $U : [0, \infty) \to \mathbf{R}$ with $U(0) = 0$, define

$$\tilde{U}(r) = \frac{U(V(A) r)}{V(A)}. \qquad (91)$$

Then $\tilde{U}'(\infty) = U'(\infty)$ and $U \in DC_N$ if and only if $\tilde{U} \in DC_N$ .Now

$$U_{v'}(\mu)=\int_A U(\rho')d\mathrm{V}'+U'(\infty)\mu_s(A)$$

$$=\frac{1}{\mathrm{V}(A)}\int_A U(\mathrm{V}(A)\rho)\,d\mathrm{V}+U'(\infty)\mu_s(A)$$

$$=\int_X \tilde{U}(\rho)\,d\mathrm{V}+\tilde{U}'(\infty)\mu_s(X)=\tilde{U}_v(\mu). \qquad (92)$$

As $\mathrm{P}_2(A,\mathrm{V}')\subset \mathrm{P}_2(X,\mathrm{V})$, part (a) follows.

Letting $\tilde{p}$ denote the pressure of $\tilde{U}$ one fine that

$$\frac{\tilde{p}(r)}{r}=\frac{\tilde{p}(\mathrm{V}(A)r)}{\mathrm{V}(A)r}. \qquad (93)$$

Then with reference to Definition (4.2.6) $\lambda(\tilde{U})=\lambda(U)$. Part (b) follows.

We considered the following result.

Theorem (4.2.8)[121]: Let $\{(X_i,d_i,\mathrm{V}_i)\}_{i=1}^{\infty}$ be a sequence of compact measured length spaces with $\lim_{i\to\infty}(X_i,d_i,\mathrm{V}_i)=(X,d,\mathrm{V})$ in the measured Gromov-Hausdorff topology.

If each (X_i,d_i,V_i) has nonnegative N-Ricci curvature then (X,d,V) has nonnegative N-Ricci curvature. If each (X_i,d_i,V_i) has ∞-Ricci curvature bounded below by K, for some $K\in\mathbf{R}$, then (X,d,V) has ∞-Ricci curvature bounded below by K.

We first show that a weak displacement convexity assumption implies that the measure V either is a delta function or is nonatomic.

Proposition (4.2.9)[121]: Let (X,d,V) be a compact measured length space. For all $N\in(1,\infty]$, if $H_{N,v}$ is weakly λ-displacement convex then V either is a delta function or is nonatomic.

Proof: We will assume that $\mathrm{V}(\{x\})\in(0,1)$ for some $x\in X$ and derive a contradiction.

Suppose first that $N\in(1,\infty)$. Put $\mu_0=\delta_x$ and $\mu_1=\frac{\mathrm{V}-\mathrm{V}(\{x\})\delta_x}{1-\mathrm{V}(\{x\})}$. By the hypothesis and Proposition (4.1.10): there is a displacement interpolation $\{\mu_t\}_{t\in[0,1]}$ from μ_0 to μ_1 along which

$U_v(\mu_t) \le tU_v(\mu_1) + (1-t)U_v(\mu_0) - \frac{1}{2}\lambda t(1-t)\mathrm{W}_2(\mu_0,\mu_1)^2$ Satisfied with $U_{\mathrm{V}} = H_{N,\mathrm{V}}$.. Now

$H_{N,\mathrm{V}}(\mu_0) = N - N\left(\mathrm{V}(\{x\})\right)^{1/N}$ and $H_{N,\mathrm{V}}(\mu_1) = N - N\left(1-\mathrm{V}(\{x\})\right)^{1/N}$. Hence

$$H_{N,\mathrm{V}}(\mu_t) \le N - (1-t)N\left(\mathrm{V}(\{x\})\right)^{1/N} - tN\left(1-\mathrm{V}(\{x\})\right)^{1/N} - \frac{1}{2}\lambda t(1-t)\mathrm{W}_2(\mu_0,\mu_1)^2. \quad (94)$$

Put $D = diam(X)$.. As we have a displacement interpolation, it follows that if $t>0$ then $\mathrm{supp}(\mu_t) \subset \overline{B_{tD}(x)}$ and $\mu_t(\{x\}) = 0$. Letting $\mu_t = \rho_t v + (\mu_t)_s$ be the Lebesgue decomposition of μ_t with respect to V, Holder's inequality implies that

$$\int_x \rho_t^{1-\frac{1}{N}} = \int_{\overline{B_{tD}(x)}-\{x\}} \rho_t^{1-\frac{1}{N}} d\mathrm{V}$$

$$\le \left(\int_{\overline{B_{tD}(x)}-\{x\}} \rho_t d\mathrm{V}\right)^{1-\frac{1}{N}} \mathrm{V}\left(\overline{B_{tD}(x)} - \{x\}\right)^{\frac{1}{N}} \le \mathrm{V}\left(\overline{B_{tD}(x)} - \{x\}\right)^{\frac{1}{N}}. \quad (95)$$

Then

$$H_{N,\mathrm{V}}(\mu_t) \ge N - N\left(\mathrm{V}\left(\overline{B_{tD}(x)}\right) - \mathrm{V}(\{x\})\right)^{\frac{1}{N}}. \quad (96)$$

As $\lim_{t\to 0+}\mathrm{V}\left(\overline{B_{tD}(x)}\right) = \mathrm{V}(\{x\})$, we obtain a contradiction with (94) when t is small.

If $N=\infty$ then $H_{\infty,\mathrm{V}}(\mu_0) = \log\frac{1}{\mathrm{V}(\{x\})}$ and $H_{\infty,\mathrm{V}}(\mu_1) = \log\frac{1}{1-\mathrm{V}(\{x\})}$. Hence

$$H_{\infty,\mathrm{V}}(\mu_t) \le (1-t)\log\frac{1}{\mathrm{V}(\{x\})} + t\log\frac{1}{1-\mathrm{V}(\{x\})} - \frac{1}{2}\lambda t(1-t)\mathrm{W}_2(\mu_0,\mu_1)^2. \quad (97)$$

In particular, μ_t is absolutely continuous with respect to V. Write $\mu_t = \rho_t\mathrm{V}$.. Jensen's

inequality implies that for $t>0$,

$$\int_{\overline{B_{tD}(x)}-\{x\}} \rho_t \log(\rho_t) \frac{dV}{V\left(\overline{B_{tD}(x)}-\{x\}\right)} \geq$$

$$\left(\int_{\overline{B_{tD}(x)}-\{x\}} \rho_t \frac{dV}{V\left(\overline{B_{tD}(x)}-\{x\}\right)}\right).\log\left(\int_{\overline{B_{tD}(x)}-\{x\}} \rho_t \frac{dV}{V\left(\overline{B_{tD}(x)}-\{x\}\right)}\right)$$

$$= \frac{1}{V\left(\overline{B_{tD}(x)}-\{x\}\right)} \log\left(\frac{1}{V\left(\overline{B_{tD}(x)}-\{x\}\right)}\right). \tag{98}$$

Then

$$H_{\infty,V}(\mu_t) = \int_X \rho_t \log(\rho_t)\, dV = \int_{\overline{B_{tD}(x)}-\{x\}} \rho_t \log(\rho_t)\, dV$$

$$\geq \log\left(\frac{1}{V\left(\overline{B_{tD}(x)}-\{x\}\right)}\right). \tag{99}$$

As $\lim_{t\to 0+V}\left(B_{tD}(x)\right) = V(\{x\})$, we obtain a contradiction with (97) when t is small.

We now prove a Bishop-Gromov-type inequality.

Proposition (4.2.10)[121]: Let (X,d,V) be a compact measured length space. Assume that $H_{N,V}$ is weakly displacement convex on $P_2(X)$, for some $N \in (1,\infty)$. Then for all $x \in \operatorname{supp}(V)$ and all $0 < r_1 \leq r_2$,

$$V\left(B_{r_2}(x)\right) \leq \left(\frac{r_2}{r_1}\right)^N V\left(B_{r_1}(x)\right). \tag{100}$$

Proof: From Proposition (4.2.9): we may assume that V is nonatomic, as the theorem is trivially true when $V = \delta_x$. Put $\mu_0 = \delta_x$ and $\mu_1 = \frac{1_{B_{r_2}(x)}}{V\left(B_{r_2}(x)\right)} V.$. By the hypothesis and Proposition (4.1.10): there is a displacement interpolation $\{\mu_t\}_{t\in[0,1]}$ from μ_0 to μ_1 along which $U_v(\mu_t) \leq tU_v(\mu_1) + (1-t)U_v(\mu_0) - \frac{1}{2}\lambda t(1-t) W_2(\mu_0,\mu_1)^2$ is satisfied with $U_v = H_{N,v}$ and $\lambda = 0.$. Now $H_{N,v}(\mu_0) = N$ and $H_{N,V}(\mu_1) = N - N\left(V\left(B_{r_2}(x)\right)\right)^{1/N}$.. Hence

$$H_{N,V}(\mu_t) \le N - tN\left(V\left(B_{r_2}(x)\right)\right)^{1/N}. \quad (101)$$

Let $\mu_t = \rho_t V + (\mu_t)_s$ be the Lebesgue decomposition of μ_t with respect to V. As we have a displacement interpolation, ρ_t vanishes outside of $B_{t\,r_2}(x)$. Then from Holder's inequality,

$$H_{N,V}(\mu_t) \ge N - N\left(V\left(B_{tr_2}(x)\right)\right)^{1/N}. \quad (102)$$

The theorem follow by taking $t = \frac{r_1}{r_2}$.

Theorem (4.2.11)[121]: If a compact measured length space (X,d,V) has nonnegative N–Ricci curvature for some $N \in [1,\infty)$ then for all $x \in \text{supp}(V)$ and all $0 < r_1 \le r_2$,

$$V\left(B_{r_2}(x)\right) \le \left(\frac{r_2}{r_1}\right)^N V\left(B_{r_1}(x)\right). \quad (103)$$

Corollary (4.2.12) [121]: Given $N \in [1,\infty)$ and $D \ge 0$, the space of compact measured length spaces (X,d,V) with nonnegative N-Ricci curvature, $diam(X,d) \le D$ and $\text{supp}(V) = X$ is sequentially compact in the measured Gromov-Hausdorff topology.

Proof: Let $\{(X_i,d_i,V_i)\}_{i=1}^{\infty}$ be a sequence of such spaces. Using the Bishop-Gromov inequality of Theorem (4.2.11): along with the fact that $\text{supp}(V_i) = X_i$, it follows as in [166] that after passing to a subsequence we may assume that $\{(X_i,d_i)\}_{i=1}^{\infty}$ converges in the Gromov-Hausdorff topology to a compact length space (X,d), necessarily with $diam(X,d) \le D$. Let $f_i : X_i \to X$ be Borel ε_i-approximations, with $\lim_{i\to\infty} \varepsilon_i = 0.$. From the compactness of $P_2(X)$,, after passing to a subsequence we may assume that $\lim_{i\to\infty} (f_i)_* V_i = V.$ for some $V \in P_2(X)$. From Theorem (4.2.8): (X,d,V) has nonnegative N-Ricci curvature.

It remains to show that $\text{supp}(V) = X$. Given $x \in X$, the measured Gromov Hausdorff convergence of $(X_i,d_i,V_i)_{i=1}^{\infty}$ to (X,d,V) implies that there is a sequence of points $x_i \in X_i$ with $\lim_{i\to\infty} f_i(x_i) = x$ so that for all $r > 0$ and

$\varepsilon \in (0,r)$, we have $\limsup_{i\to\infty} \overline{V_i(B_{r-\varepsilon}(x_i))} \le V(B_r(x))$. By Theorem (4.2.11): $(r-\varepsilon)^{-N} V_i(B_{r-\varepsilon}(x_i)) \ge daim(X,d)^{-N}$ Then $V(B_r(x)) \ge \left(\frac{r}{diam(X,d)}\right)^N$, which proves the claim.

We show that in certain cases, lower Ricci curvature bounds are preserved upon quotienting by a compact group action.

Lemma (4.2.13)[121]: The map $p_*: P_2(X) \to P_2(X/G)$ restricts to an isometric isomorphism between the set $P_2(X)^G$ of G-invariant elements in $P_2(X)$, and $P_2(X/G)$.

Proof: Let dh be the normalized Haar measure on G. The map $p_*: P_2(X) \to P_2(X/G)$ restricts to an isomorphism $p_*: P_2(X)^G \to P_2(X/G)$; the problem is to show that it is an isometry. Let $\tilde{\pi}$ be a transference plan between $\tilde{\mu}_0, \tilde{\mu}_1 \in P_2(X)^G$ Then $\tilde{\pi}' = \int_G g.\tilde{\pi} dh(g)$ is also a transference plan between $\tilde{\mu}_0$ and $\tilde{\mu}_1$ with

$$\int_{X\times X} dx(\tilde{x},\tilde{y})^2 d\tilde{\pi}'(\tilde{x},\tilde{y}) = \int_G \int_{X\times X} dx(\tilde{x}g,\tilde{y}g)^2 d\tilde{\pi}'(\tilde{x},\tilde{y})\, dh(g) = \int_{X\times X} dx(\tilde{x},\tilde{y})^2 d\pi(\tilde{x},\tilde{y}) \quad (104)$$

Thus there is a G-invariant optimal transference plan $\tilde{\pi}$ between $\tilde{\mu}_0$ and $\tilde{\mu}_1$. As $\pi = (p\times p)_* \tilde{\pi}$ is a transference plan between $p_*\tilde{\mu}_0$ and $p_*\tilde{\mu}_1$, with

$$\int_{(X/G)\times(X/G)} d_{x/G}(x,y)^2 d\pi(x,y) = \int_{X\times X} d_{x/G}(p(\tilde{x}),p(\tilde{y}))^2 d\tilde{\pi}(\tilde{x},\tilde{y}) \le \int_{X\times X} d_X(\tilde{x},\tilde{y})^2 d\tilde{\pi}(\tilde{x},\tilde{y}), \quad (105)$$

it follows that the map $p_*: P_2(X)^G \to P_2(X/G)$ is metrically nonincreasing.

Conversely, let $s:(X/G)\times(X/G) \to X\times X$ be a Borel map such that $(p\times p)\circ s = Id$ and $d_X \circ s = d_{X/G}$ That is, given $x,y \in X/G$, the map s picks points $\tilde{x} \in p^{-1}(x)$ and $\tilde{y} \in p^{-1}(y)$ in the corresponding orbits so that the distance between $\tilde{x}$ and $\tilde{y}$ is minimized among all pairs of points in $p^{-1}(x)\times p^{-1}(y)$. (The existence of s follows from applying [240] to the restriction of $p\times p$ to $\{(\tilde{x},\tilde{y}) \in X\times X : d_X(\tilde{x},\tilde{y}) = d_{X/G}(p(\tilde{x}),p(\tilde{y}))\}$. The restriction map is a surjective Borel map with compact preimages.) Given an optimal transference plan π between $\mu_0, \mu_1 \in P_2(X/G)$,, define a measure π on $X\times X$ by saying that for all $\tilde{F} \in C(X\times X)$,

$$\int_{X\times X}\tilde{F}d\tilde{\pi}=\int_G\int_{(X/G)\times(X/G)}\tilde{F}\left(s(x,y).(g,g)\right)dx(x,y)dh(g). \tag{106}$$

Then for $F\in C\left((X/G)\times(X/G)\right)$,

$$\int_{(X/G)\times(X/G)}Fd\left(p\times p\right)_*\tilde{\pi}=\int_{X\times X}Fd\left(p\times p\right)^*d\tilde{\pi}$$

$$=\int_G\int_{(X/G)\times(X/G)}\left(p\times p\right)^*F\left(s(x,y).(g,g)\right)d\pi(x,y)dh(g)$$

$$=\int_G\int_{(X/G)\times(X/G)}\left(p\times p\right)\left(s(x,y).(g,g)\right)d\pi(x,y)dh(g)$$

$$=\int_{(X/G)\times(X/G)}F(x,y)\,d\pi(x,y). \tag{107}$$

Thus $(p\times p)_*\tilde{\pi}=\pi$. As $\tilde{\pi}$ is G-invariant, it follows that it is a transference plan between $(p_*)^{-1}(\mu_0),(p_*)^{-1}(\mu_1)\in P_2(X)^G$.Now

$$\int_{X\times X}d_X(\tilde{x},\tilde{y})^2\,d\tilde{\pi}(\tilde{x},\tilde{y})=\int_G\int_{(X/G)\times(X/G)}d_X\left(s(x,y)(g,g)\right)^2d\pi(x,y)dh(g)$$

$$=\int_{(X/G)\times(X/G)}d_{X/G}(x,y)^2\,d\pi(x,y). \tag{108}$$

Thus p_* and $(p_*)^{-1}$ are metrically non-increasing, which shows that p_* defines an isometric isomorphism between $P_2(X)^G$ and $P_2(X/G)$..

Theorem (4.2.14)[121]: Let (X,d,V) be a compact measured length space. Suppose that any two $\mu_0,\mu_1\in P_2^{ac}(X,V)$ are joined by a unique Wasserstein geodesic, that lies in $P_2^{ac}(X,V)$. Suppose that a compact topological group G acts continuously and isometrically on X preserving V. Let $p:X\to X/G$ be the quotient map and let $d_{X/G}$ be the quotient metric. We have the following implications

(a) For $N\in[1,\infty)$, if (X,d,V) has nonnegative N-Ricci curvature then $(X/G,d_{X/G},P_*V)$ has nonnegative N- Ricci curvature.

(b) If (X,d,V) has ∞-Ricci curvature bounded below by K then $(X/G,d_{X/G},P_*V)$ has ∞ – Ricci curvature bounded below by K.

Proof : The proofs of parts a. and b. of the theorem are similar, so we will be content with proving just part(a).First, $(X/G, d_{X/G})$is a length space. (Given $x, y \in X/G$, let $\tilde{x} \in p^{-1}(x)$ and $\tilde{y} \in p^{-1}(y)$ satisfy $d_X(\tilde{x}, \tilde{y}) = d_{X/G}(x, y)$. If c is a geodesic from $\tilde{x}$ to $\tilde{y}$ then $p \circ c$ is a geodesic from x to y.)

Given $\mu_0, \mu_1 \in \mathrm{P}_2^{ac}(X/G, p_* \mathrm{V})$write $\mu_0 = \rho_0 p_* \mathrm{V}$ and $\mu = \rho_1 p_* \mathrm{V}$.Put $\tilde{\mu}_0 = (p^* \rho_0)\mathrm{V}$ and $\tilde{\mu}_t = (p^* \rho_1)\mathrm{V}$.. From Lemma (4.2.13): $\mathrm{W}_2(\tilde{\mu}_0, \tilde{\mu}_1) = \mathrm{W}^2(\mu_0, \mu_1)$. By hypothesis, there is a Wasserstein geodesic $\{\tilde{\mu}_t\}_{t\in[0,1]}$from $\tilde{\mu}_0$ to $\tilde{\mu}_1$ so that for all $U \in \mathrm{DC}_N$, equation $U_v(\mu_t) \le tU_v(\mu_1) + (1-t)U_v(\mu_0) - \frac{1}{2}\lambda t(1-t)\mathrm{W}_2(\mu_0, \mu_1)^2$ in section is satisfied along $\{\tilde{\mu}_t\}_{t\in[0,1]}$, with $\lambda = 0$. The geodesic $\{\tilde{\mu}_t\}_{t\in[0,1]}$ is G-invariant, as otherwise by applying an appropriate element of G we would obtain two distinct Wasserstein geodesics between $\tilde{\mu}_0$ and $\tilde{\mu}_1$. Put $\mu_t = p_* \tilde{\mu}_t$. It follows from the above discussion that $\{\tilde{\mu}_t\}_{t\in[0,1]}$ is a curve with length $\mathrm{W}_2(\tilde{\mu}_0, \tilde{\mu}_1)$, and so is a Wasserstein geodesic. As $\tilde{\mu}_t \in \mathrm{P}_2^{ac}(X, \mathrm{V})$, we have $\tilde{\mu}_t \in \mathrm{P}_2^{ac}(X/G, p_* \mathrm{V})$ Write $\mu_t = \rho_t p_* \mathrm{V}$..Then $\tilde{\mu}_t = (p^* \rho_t)\mathrm{V}$. As

$$U_{p,\mathrm{V}}(\mu_t) = \int_{X/G} U(\rho_t)\, dp_* \mathrm{V} = \int_X p^* U(\rho_t)\, d\mathrm{V} = \int_X U(p^* \rho_t)\, d\mathrm{V} = U_v(\tilde{\mu}_t). \qquad (109)$$

it follows that equation$U_v(\mu_t) \le tU_v(\mu_1) + (1-t)U_v(\mu_0) - \frac{1}{2}\lambda t(1-t)\mathrm{W}_2(\mu_0, \mu_1)^2$ in section is satisfied along $\{\tilde{\mu}_t\}_{t\in[0,1]}$with $\lambda = 0$.Along with proposition (4.1.15): this concludes the proof of part (a).

Lemma (4.2.15)[121]: Let $\{\mu_i\}_{i=1}$ be a finite subset of $\mathrm{P}_2^{ac}(X, \mathrm{V})$, with densities $\rho_i = \frac{d\mu_i}{d\mathrm{V}}$.. If $N < \infty$ then there is a function $U \in \mathrm{DC}_N$ such that

$$\lim_{r\to\infty} \frac{U(r)}{r} = \infty \qquad (110)$$

and

$$\sup_{1\le i\le m} \int_X U(\rho_i(x))\, d\mathrm{V}(x) < \infty. \qquad (111)$$

Proof: As a special case of the Dunford-Pettis theorem [81], there is an increasing function $\Phi : (0, \infty) \to \mathbf{R}$ such that

$$\lim_{r\to\infty}\frac{\Phi(r)}{r}=\infty \tag{112}$$

and

$$\sup_{1\le i\le m}\int_X \Phi(\rho_i(x))d\mathrm{V}(x)<\infty. \tag{113}$$

We may assume that Φ is identically zero on [0,1].

Consider the function $\phi:(0,\infty)\to\mathbf{R}$ given by

$$\phi(\lambda)=\lambda^N\Phi(\lambda^{-N}). \tag{114}$$

Then $\phi\equiv 0$ on $[1,\infty)$, and $\lim_{\lambda\to 0^+}\phi(\lambda)=\infty$. Let $\tilde{\phi}$; be the lower convex hull of ϕ; on $(0,\infty)$, i.e. the supremum of the linear functions bounded above by ϕ. Then $\tilde{\phi}\equiv 0$ on $[1,\infty)$ and $\tilde{\phi}$ is nonincreasing. We claim that $\lim_{\lambda\to 0^+}\tilde{\phi}(\lambda)=\infty$. If not, suppose that $\lim_{\lambda\to 0^+}\tilde{\phi}(\lambda)=M<\infty$. Let $a=\sup_{\lambda\ge 0}\frac{M+1-\phi(\lambda)}{\lambda}<\infty$ (because this quantity is ≤ 0 when λ is small enough). Then $\phi(\lambda)\ge M+1-a\lambda$,so $\lim_{\lambda\to 0^+}\tilde{\phi}(\lambda)\ge M+1,$ which is a contradiction.

Now set

$$U(r)=r\tilde{\phi}(r^{-1/N}). \tag{115}$$

Since $\tilde{\phi}\le\phi$; and $\Phi(r)=r\phi(r^{-1/N})$, we see that $U\le\Phi.$ Hence

$$\sup_{1\le i\le m}\int_X U(\rho_i(x))\; dv(x)<\infty. \tag{116}$$

Since $\lim_{\lambda\to 0^+}\tilde{\phi}(\lambda)=\infty$ we also know that

$$\lim_{r\to\infty}\frac{U(r)}{r}=\infty . \tag{117}$$

Clearly U is continuous with $U(0)=0.$ As $\tilde{\phi}$ is convex and nonincreasing, it follows that U is convex. Hence $U\in \mathrm{DC}_N$.

Theorem(4.2.16)[121]: If (X,d,V) has nonnegative N-Ricci curvature for some $N\in[1.\infty)$ then $P_2^{ac}(X,\mathrm{V})$ is a convex subset of $P_2(X)$..

Proof: Given $\mu_0,\mu_1 \in \mathrm{P}_2^{ac}(X,\mathrm{V})$ put $\rho_0 = \frac{d\mu_0}{dv}$ and $\rho_1 = \frac{d\mu}{dv}$ By Lemma (4.2.15) there is a $U \in \mathrm{DC}_N$. with $U'(\infty)=\infty$ such that $U_v(\mu_0)<\infty$ and $U_v(\mu_1)<\infty$.As (X,d,V) has nonnegative N-Ricci curvature, there is a Wasserstein geodesic $\{\mu_t\}_{t\in[0,1]}$ from μ_0 to μ_1 so that (117) is satisfied with $\lambda=0$.In particular, $U_v(\mu_t)<\infty$ for all $t\in[0,1]$. As $U'(\infty)=\infty$, it follows that $\mu_t \in \mathrm{P}_2^{ac}(X,\mathrm{V})$ for each t.

We now clarify the relationship between (X,d,V) having nonnegative N-Ricci curvature and the analogous statement for $\mathrm{supp}(\mathrm{V})$.

Theorem (4.2.17)[121]: (a) Given $N\in[1,\infty)$, suppose that a compact measured length space (X,d,V) has nonnegative N -Ricci curvature. Then $\mathrm{supp}(\mathrm{V})$ is a convex subset of X (although not necessarily totally convex) and $\left(\mathrm{supp}(\mathrm{V}),d\big|_{\mathrm{supp}(\mathrm{V})},\mathrm{V}\right)$ has nonnegative

N-Ricci curvature. Conversely, if $\mathrm{supp}(\mathrm{V})$ is a convex subset of X and $\left(\mathrm{supp}(\mathrm{V}),d\big|_{\mathrm{supp}(\mathrm{V})},\mathrm{V}\right)$ has nonnegative $N-$ Ricci curvature then (X,d,V)has nonnegative N- Ricci curvature.

(b) Given $K\in\mathbf{R}$ the analogous statement holds when one replaces "nonnegative N-Ricci curvature" by "∞-Ricci curvature bounded below by K".

Proof: (a) Let (X,d,V)be a compact measured length space with nonnegative N-Ricci curvature. Let μ_0 and μ_1 be elements of $\mathrm{P}_2(X,\mathrm{V})$ There are sequences $\{\mu_{k,0}\}_{k=1}^{\infty}$ and $\{\mu_{k,1}\}_{k=1}^{\infty}$ in $\mathrm{P}_2^{ac}(X,\mathrm{V})$(in fact with continuous densities) such that $\lim_{k\to\infty}\mu_{k,0}=\mu_0 \lim_{k\to\infty}\mu_{k,1}=\mu_1$ and for all $U\in\mathrm{DC}_N \lim_{k\to\infty}U_v(\mu_{k,0})=U_v(\mu_0)$ and $\lim_{k\to\infty}U_{\mathrm{V}}(\mu_{k,1})=U_{\mathrm{V}}(\mu_1)$. From the definition of nonnegative N-Ricci, for each k there is a Wasserstein geodesic $\{\mu_{k,t}\}_{t\in[0,1]}$ such that

$$U_{\mathrm{V}}(\mu_{k,t})\le tU_{\mathrm{V}}(\mu_{k,t})+(1-t)U_{\mathrm{V}}(\mu_{k,0}) \tag{118}$$

for all $U\in\mathrm{DC}_N$ and $t\in[0,1]$ By repeating the proof of Theorem(4.2.16) each $\mu_{k,t}$ is absolutely continuous with respect to V. In particular, it is supported in $\mathrm{supp}(\mathrm{V})$. By the same reasoning as in the proof of Proposition (4.1.15) after passing to a subsequence we may assume that as $K\to\infty$, the geodesics $\{\mu_{k,t}\}_{t\in[0,1]}$ converge uniformly to a Wasserstein geodesic $\{\mu_{k,t}\}_{t\in[0,1]}$ that satisfies

$$U_{\mathrm{V}}(\mu_t) \le tU_{\mathrm{V}}(\mu_1) + (1-t)U_{\mathrm{V}}(\mu_0). \qquad (119)$$

For each $t \in [0,1]$, the measure μ_t is the weak-$*$ limit of the probability measures $\{\mu_{k,t}\}_{k=1}^{\infty}$ which are all supported in the closed set $\mathrm{supp}(\mathrm{V})$. Hence μ_t is also supported in $\mathrm{supp}(\mathrm{V})$.

To summarize, we have shown that $\{\mu_{k,t}\}_{t\in[0,1]}$ is a Wasserstein geodesic lying in $\mathrm{P}_2(X,\mathrm{V})$ that satisfies (119) for all $U \in \mathrm{DC}_N$ and $t \in [0,1]$

We now check that $\mathrm{supp}(\mathrm{V})$ is convex. Let x_0 and x_1 be any two points in $\mathrm{supp}(\mathrm{V})$. Applying the reasoning above to $\mu_0 = \delta_{x_0}$ and $\mu_1 = \delta_{x_1}$ one obtains the existence of a Wasserstein geodesic $\{\mu_{k,t}\}_{t\in[0,1]}$ joining δ_{x_0} to δ_{x_1} such that each μ_t is supported in $\mathrm{supp}(\mathrm{V})$. By Proposition (4.2.10) there is an optimal dynamical transference plan $\Pi \in \mathrm{P}(\Gamma)$ such that $\mu_t = (e_t)_* \Pi$ for all $t \in [0,1]$. For each $t \in [0,1]$, we know that $\gamma(t) \in \mathrm{supp}(\mathrm{V})$ holds for Π almost all γ. It follows that for Π-almost all γ we have $\gamma(t) \in \mathrm{supp}(\mathrm{V})$, for all $t \in \mathrm{Q} \cap [0,1]$. As $\gamma \in \Gamma$ is continuous, this is the same as saying that for Π-almost all γ, the geodesic γ is entirely contained in $\mathrm{supp}(\mathrm{V})$. Also, for Π-almost all γ we have $\gamma(0) = x_0$ and $\gamma(1) = x_1$. Thus x_0 and x_1 are indeed joined by a geodesic path contained in $\mathrm{supp}(\mathrm{V})$.

This proves the direct implication in part a. The converse is immediate.

(b) The proof of part (b) follows the same lines as that of part a. We construct the approximants $\{\mu_{k,0}\}_{k=1}^{\infty}$ and $\{\mu_{k,t}\}_{k=1}^{\infty}$, with continuous densities, and the geodesics $\{\mu_{k,t}\}_{t\in[0,1]}$.

As $H_{\infty,v}(\mu_{0,k}) < \infty$ and $H_{\infty,v}(\mu_{1,k}) < \infty$,, we can apply inequality (117) with $U = H_\infty$

and $\lambda = K$, to deduce that $H_{\infty,v}(\mu_{t,k}) < \infty$, for all $t \in [0,1]$. This implies that $\mu_{t,k}$ is absolutely continuous with respect to V. The rest of the argument is similar to that of part (a)

Theorem (4.2.18)[121]: Suppose that (X,d,V) has ∞-Ricci curvature bounded below by $K > 0$. Then for all $\mu \in \mathrm{P}_2(X,\mathrm{V})$,,

$$\frac{K}{2} W_2(\mu, V)^2 \le H_{\infty,V}(\mu). \tag{120}$$

If now $\mu \in P_2^{ac}(X, V)$ and its density_ $\rho = \frac{d\mu}{dV}$ is a positive Lipschitz function on X then

$$H_{\infty,V}(\mu) \le W_2(\mu, V)\sqrt{1_{\infty,V}(\mu)} - \frac{K}{2} W_2(\mu, V)^2 \frac{1}{2K} 1_{\infty,V}(\mu). \tag{121}$$

If on the other hand (X, d, V) has ∞-Ricci curvature_bounded below by $K \le 0$ then

$$H_{\infty,v}(\mu) \le diam(X)\sqrt{1_{\infty,v}(\mu)} - \frac{K}{2} diam(X)^2. \tag{122}$$

If (X, d, V) has nonnegative N -Ricci curvature then

$$H_{N,v}(\mu) \le diam(X)\sqrt{1_{N,v}(\mu)}. \tag{123}$$

We now express the conclusion of Theorem(4.2.18): in terms of more standard inequalities, starting with the case $N = \infty$.

(i) The case $N = \infty$.

Definition (4.2.19)[121]: Suppose that $K > 0$.

We say that V satisfies a log Sobolev inequality with constant K, $LSI(K)$, if for all $\mu \in P_2^{ac}(X, V)$ whose density $\rho = \frac{d\mu}{dV}$ is Lipschitz and positive, we have

$$H_{\infty,V}(\mu) \le \frac{1}{2K} 1_{\infty,V}(\mu). \tag{124}$$

We say that V satisfies a Talagmnd inequality with constant $K, T(K)$, if for all $\mu \in P_2(X, V)$

$$W_2(\mu, V) \le \sqrt{\frac{2H_{\infty,V}(\mu)}{K}}. \tag{125}$$

We say that V satisfies a Poincare inequality with constant $K, P(K)$, if for all $h \in Lip(X)$ with $\int_X h dV = 0$, we have

$$\int_X h^2 d\mathrm{V} \le \frac{1}{K}\int_X \left|\nabla^- h\right|^2 d\mathrm{V}. \quad (126)$$

All of these inequalities are associated with concentration of measure [20, 256, 254, 174]. For example, $\mathrm{T}(\mathrm{K})$ implies a Gaussian-type concentration of measure.

The following chain of implications, none of which is an equivalence, is well-known in the context of smooth Riemannian manifolds:

$$[Ric \ge K] \Rightarrow LSI(K) \Rightarrow T(K) \Rightarrow \mathrm{P}\,(K). \quad (127)$$

In the context of length spaces, we see from Theorem (4.2.18) that having ∞-Ricci curvature bounded below by $K > 0$ implies $\mathrm{LSI}(\mathrm{K})$ and $\mathrm{T}(\mathrm{K})$. The next corollary makes the statement of the log Sobolev inequality more explicit.

Corollary (4.2.20)[121]: Suppose that (X, d, V) has ∞-Ricci curvature bounded below by $K \in \mathbf{R}$

If $f \in Lip(X)$ satisfies $\int_X f^2 d\mathrm{V} = 1$ then

$$\int_X f^2 \log\left(f^2\right) d\mathrm{V} \le 2\mathrm{W}_2\left(f^2\mathrm{V}, \mathrm{V}\right)\sqrt{\int_X \left|\nabla^- f\right|^2 dv} - \frac{K}{2}\mathrm{W}_2\left(f^2\mathrm{V}, \mathrm{V}\right)^2. \quad (128)$$

In particular, if $K > 0$ then

$$\int_X f^2 \log\left(f^2\right) d\mathrm{V} \le \frac{2}{K}\int_X \left|\nabla^- f\right|^2 d\mathrm{V}, \quad (129)$$

while if $K \le 0$ then

$$\int_X f^2 \log\left(f^2\right) d\mathrm{V} \le 2diam(X)\sqrt{\int_X \left|\nabla^- f\right|^2 d\mathrm{V}} - \frac{K}{2} diam(X)^2. \quad (130)$$

Proof: For any $\varepsilon \to 0$, put $\rho_t = \dfrac{f^2 + \varepsilon}{1+\varepsilon}$ From Theorem (4.2.18)

$$\int_X \rho_\varepsilon \log(\rho_\varepsilon) d\mathrm{V} \le \mathrm{W}_2(\rho_\varepsilon, \mathrm{V}, \mathrm{V})\sqrt{\int_X \frac{\left|\nabla^- f\right|^2}{\rho_\varepsilon} d\mathrm{V}} - \frac{K}{2}\mathrm{W}_2(\rho_\varepsilon, \mathrm{V}, \mathrm{V})^2 \quad (131)$$

As

$$\frac{\left|\nabla^{-}\rho_{\varepsilon}\right|^{2}}{\rho_{\varepsilon}}=\frac{1}{1+\varepsilon}\frac{4f^{2}}{f^{2}+\varepsilon}\left|\nabla^{-}f\right|^{2}, \tag{132}$$

the corollary follows by taking. $\varepsilon \to 0$.

We now recall standard fact that $\mathrm{LSI}(\mathrm{K})$ implies $\mathrm{P}(\mathrm{K})$.

Theorem(4.2.21)[121]: Let (X,d,V) be a compact measured length space satisfying $\mathrm{LSI}(\mathrm{K})$ for some $K>0$. Then it also satisfies $\mathrm{P}(\mathrm{K})$.

Proof: Suppose_that $h\varepsilon Lip(X)$ satisfies $\int_X hdv=0.$. For $\varepsilon\in[0,\frac{1}{\|h\|_\infty})$ put

$f_\varepsilon=\sqrt{1+\varepsilon h}>0.$ As $2f_\varepsilon\nabla^{-}f_\varepsilon=\varepsilon\nabla^{-}h,$ it follows that

$$\lim_{\varepsilon\to\infty^{+}}\left(\frac{1}{\varepsilon^{2}}\int_X\left|\nabla^{-}f_{\varepsilon}\right|^{2}d\mathrm{V}\right)=\frac{1}{4}\int_X\left|\nabla^{-}h\right|^{2}d\mathrm{V}. \tag{133}$$

As the Taylor expansion of $x\log(x)-x+1$ around $x=1$ is $\frac{1}{2}(x-1)^{2}+\ldots,$ it follows that

$$\lim_{\varepsilon\to\infty^{+}}\frac{1}{\varepsilon^{2}}\int_X f_{\varepsilon}^{2}\log\left(f_{\varepsilon}^{2}\right)d\mathrm{V}=\frac{1}{2}\int_X h^{2}d\mathrm{V}. \tag{134}$$

Then the conclusion follows from (129).

As mentioned above, in the case of smooth Riemannian manifolds there are stronger implications: $T(K)$ implies $\mathrm{P}(K)$, and $\mathrm{LSI}(\mathrm{K})$ implies $T(K)$. We will show elsewhere that the former is always true, while the latter is true under the additional assumption of a lower bound the Alexandrov curvature:

Theorem(4.2.22)[121]: Let (X,d,V) be a compact measured length space.

(i) If V satisfies $T(K)$ for some $K>0$, then it also satisfies $\mathrm{P}(K)$.

(ii) If X is a finite-dimensional Alexandrov space with Alexandrov curvature bounded below, and satisfies $\mathrm{LSI}(\mathrm{K})$ for some $K>0$, then it also satisfies $T(K)$

N-Ricci curvature, with $N<\infty$, then it admits a local Poincare inequality, at least if one assumes almost-everywhere uniqueness of geodesics. We will discuss this in detail elsewhere.

The case $N<\infty$. We now write an analog of Corollary(4.2.20) in the case $N<\infty$. Suppose that (X,d,V) has nonnegative N-Ricci curvature. Then if ρ is a positive Lipschitz function on X, (123) says that

$$N-N\int_X \rho^{1-\frac{1}{N}}d\mathrm{V}\leq\frac{N-1}{N}diam(X)\sqrt{\int_X\frac{\left|\nabla^-\rho\right|^2}{\rho^{\frac{2}{N}+1}}d\mathrm{V}} \tag{135}$$

If $N>2$ put $f=\rho^{\frac{N-2}{2N}}$. Then $\int_X f^{\frac{2N}{N-2}}d\mathrm{V}=1$ and one finds that (135) is equivalent to

$$1-\int_X \rho^{\frac{2(N-1)}{(N-2)}}d\mathrm{V}\leq\frac{2(N-1)}{N(N-2)}diam(X)\sqrt{\int_X\left|\nabla^- f\right|^2d\mathrm{V}}. \tag{136}$$

As in the proof of Corollary (4.2.20) equation (136) holds for all $f\in Lip(X)$ satisfying $\int_X f^{\frac{2N}{N-2}}d\mathrm{V}=1$. From $H\ddot{O}lder's$ inequality

$$\int_X f^{\frac{2(N-1)}{N-2}}d\mathrm{V}\leq\left(\int_X f\,d\mathrm{V}\right)^{\frac{2}{N+2}}\left(\int_X f^{\frac{2N}{N-2}}d\mathrm{V}\right)^{\frac{N}{N+2}}=\left(\int_X fd\mathrm{V}\right)^{\frac{2}{N+2}}. \tag{137}$$

Then (136) implies

$$1\leq\frac{2(N-1)}{N(N-2)}diam(X)\sqrt{\int_X\left|\nabla^- f\right|^2d\mathrm{V}}+\left(\int_X fd\mathrm{V}\right)^{\frac{2}{N+2}}. \tag{138}$$

Writing (138) in a homogeneous form, one sees that its content is as follows: for a function F on X, bounds on $\left\|\nabla^- F\right\|_2$ and $\|F\|_1$ imply a bound on $\|F\|_{\frac{2N}{N-2}}$. This is of course an instance of Sobolev embedding

If $N=2$, putting $f=\log\left(\frac{1}{\rho}\right)$, one finds that $\int_X e^{-f}d\mathrm{V}=1$

$$1-\int_X e^{-\frac{f}{2}}d\mathrm{V}\leq\frac{1}{4}diam(X)\sqrt{\int_X\left|\nabla^- f\right|^2d\mathrm{V}}. \tag{139}$$

The classical Bonnet-Myers theorem says that if M is a smooth connected complete

N-dimensional Riemannian manifold with $Ric_M \geq K_{gM} > 0$, then $\operatorname{diam}(M) \leq \pi\sqrt{\frac{N-1}{K}}$.

Theorem(4.2.23)[121]: There is a constant $C > 0$ with the following property. Let (X, d, V) be a compact measured length space with nonnegative N-Ricci curvature, and ∞-Ricci curvature bounded below by $K > 0$. Suppose that $\operatorname{supp}(\mathrm{V}) = X$. Then

$$diam(X) \leq C\sqrt{\frac{N}{K}}. \quad (140)$$

Proof: From Theorem (4.2.11) V satisfies the growth estimate

$$\frac{\mathrm{V}(B_r(x))}{\mathrm{V}(B_{\alpha r}(x))} \leq \alpha^{-N}, \quad 0 < \alpha \leq 1. \quad (141)$$

From Theorem (4.2.18), V satisfies $T(K)$. The result follows by repeating verbatim the proof of Theorem(2.1.26): with $R = 0$, $n = N$ and $\rho = K$. □

Let *(M, g)* be a smooth compact connected n-dimensional Riemannian manifold. Let Ric denote its Ricci tensor.

Given $\phi \in C^\infty(M)$ with $\int_M e^{-\phi} dovl_M = 1$, put $d\mathrm{V} = e^{\phi} dovl_M$.

Definition (4.2.24)[121]: For $N \in [1, \infty]$, , the N-Ricci tensor of $(M, \mathrm{g}, \mathrm{V})$ is

$$Ric_N = \begin{cases} Ric + Hess(\phi) & if\ N = \infty, \\ Ric + Hess(\phi) - \frac{1}{N-1} d\psi \otimes d\phi & if\ n < N < \infty, \\ Ric + Hess(\phi) - \infty(d\psi \otimes d\phi) & if N = n, \\ -\infty & if\ N < n, \end{cases} \quad (142)$$

where by convention $\infty.0 = 0$.

The expression for Ric_∞ is the Bakry-Emery tensor [27]. The expression for Ric_N with $N < n < \infty$ was considered in [119, 273]. The statement $Ric_N \geq Kg$ is equivalent to the statement that the operator $L = \Delta - (\nabla\phi).\nabla$ satisfies Bakry's curvature-dimension condition $\mathrm{CD}(K, N)$ [25].

Given $K \in \mathbf{R}$ we recall the definition of $\lambda : \mathrm{DC}_\infty \to \mathbf{R} \cup \{-\infty\}$ from Definition

(4.2.5)

Lemma (4.2.25)[121]: Let $\phi : M \to \mathbf{R}$ be $\frac{d^2}{2}$concave function. We recall that ϕ is necessarily Lipschitz and hence $(\nabla\phi)(y)$ exists for almost all $y \in M$. For such y, define

$$F_t(y) \equiv \exp_y\left(-t\nabla\phi(y)\right). \tag{143}$$

Assume furthermore that $y \in M$ is such that

(i) ϕ admits a Hessian at y (in the sense of Alexandrov),

(ii) F_t is differentiable at y for all $t \in [0,1)$ and

(iii) $dF_t(y)$ is nonsingular for all $t \in [0,1)$.

Then $D(t) \equiv \det^{\frac{1}{n}}\left(dF_t(y)\right)$ satisfies the differential inequality

$$\frac{D''(t)}{D(t)} \le \frac{1}{n} Ric\left(F_t'(y)\right),\left(F_t'(y)\right) \qquad t \in (0,1). \tag{144}$$

Proof: Let $\{e_i\}_{i=1}^n$ be an orthonormal basis of T_yM For each i, let $J_i(t)$be defined by

$$J_i(t) = (dF_t)_y(e_i). \tag{145}$$

Then $J_i(t)_{i=1}^n$ is a Jacobi field with $J_i(0) = e_i$. Next, we note that $d\phi$ is differentiable at y, and that $d(d\phi)_y$ coincides with$Hess_y(\phi)$, up to identification. This is not so obvious (indeed, the existence of a Hessian only means the existence of a second-order Taylor expansion) but can be shown as a consequence of the semiconcavity of ϕ, as in [31]. (The case of a convex function in $\mathbf{R}^n$ is treated in [77].) It follows that

$$J_i'(0) = -Hess(\phi)(y)e_i, \tag{146}$$

Let now $\mathrm{W}(t)$ be the $n \times n$-matrix with

$$\mathrm{W}_{ij}(t) = \langle J_i(t), J_i(t)\rangle; \tag{147}$$

then $\det^{\frac{1}{n}}(dF_t)(y) = \det^{\frac{1}{2n}}\mathrm{W}(t)$.

Since W(t) is nonsingular for $t \in [0,1)$, $J_i(t)_{i=1}^n$ is a basis of $T_{F_t(y)}M$. Define a matrix $R(t)$by $J_i'(t) = \sum_j R(t)_i^j J_j(t)$It follows from the equation

$$\frac{d}{dt}\left(\left\langle J_i'(t), J_i(t) - J_i(t), J_i'(t)\right\rangle\right) = 0 \tag{148}$$

and the self-adjointness of Hess $\phi(y)$ that $\mathbf{R}\mathrm{W} - \mathbf{R}\mathrm{W}^T = 0$ for all $t \in [0,1)$, or equivalently, $\mathbf{R} = \mathrm{W}\mathbf{R}^T\mathrm{W}^{-1}$. (More intrinsically, the linear operator on $T_{F_i(y)}M$ defined by $\mathbf{R}$ satisfies $\mathbf{R} = \mathbf{R}^*$, where $\mathbf{R}^*$ is the dual defined using the inner product on $T_{F_i(y)}M$

Next,

$$\mathrm{W}' = \mathbf{R}\mathrm{W} + \mathrm{W}\mathbf{R}^T. \tag{149}$$

Applying the Jacobi equation to

$$\mathrm{W}_{ij}^n = \left\langle J_i''(t), J_j(t) + J_i(t), J_j''(t) + 2J_i'(t), J_j'(t)\right\rangle \tag{150}$$

gives

$$\mathrm{W}^n = -2Riem\left(.,F_t'(y),.,F_t'(y)\right) + 2\mathbf{R}\mathrm{W}\mathbf{R}^T. \tag{151}$$

Now

$$\frac{d}{dt}\det^{\frac{1}{2n}}\mathrm{W}(t) = \frac{1}{2n}\det^{\frac{1}{2n}}\mathrm{W}(t)Tr\left(\mathrm{W}'\mathrm{W}^{-1}\right) \tag{152}$$

and

$$\frac{d^2}{dt^2}\det^{\frac{1}{2n}}\mathrm{W}(t) = \frac{1}{4n^2}\det^{\frac{1}{2n}}\mathrm{W}(t)\left(Tr\left(\mathrm{W}'\mathrm{W}^{-1}\right)\right)^2 - \frac{1}{2n}\det^{\frac{1}{2n}}\mathrm{W}(t)Tr\left(\mathrm{W}'\mathrm{W}^{-1}\right) +$$

$$\frac{1}{2n}\det^{\frac{1}{2n}}\mathrm{W}(t)Tr\left(\mathrm{W}''\mathrm{W}^{-1}\right). \tag{153}$$

Then by (150) and(152),

$$D^{-1}\frac{d^2D}{dt^2} = \frac{1}{n^2}\left(Tr(R)\right)^2 - \frac{2}{n}Tr\left(R^2\right) - \frac{1}{n}Ric\left(F_t'(y), F_t'(y)\right) + \frac{1}{n}Tr\left(R^2\right). \tag{154}$$

As $\mathbf{R}$ is self-adjoint,

$$\frac{1}{n}\left(Tr\left(\mathbf{R}\right)\right)^2 - Tr\left(\mathbf{R}^2\right) \le 0, \qquad (155)$$

from which the conclusion follows.

Theorem (4.2.26)[121]: *a*. For $N \in (1,\infty)$, the following are equivalent.

(i) $Ric_N \ge 0$.

(ii) The measured length space $(M, \mathrm{g}, \mathrm{V})$ has nonnegative N-Ricci curvature.

(iii) For all $U \in \mathrm{DC}_N, U_v$ is weakly displacement convex on $\mathrm{P}_2(M)$.

(iv) For all $U \in \mathrm{DC}_N, U_v$ U_v is weakly a.c. displacement convex on $\mathrm{P}_2^{ac}(M)$.

(v) $H_{N,\mathrm{V}}$ is weakly a.c. displacement convex on $\mathrm{P}_2^{ac}(M)$.

b. For any $K \in \mathbf{R}$, the following are equivalent ..

(i) $Ric_N \ge Kg$.

(ii) The measured length space $(\mathrm{M}, \mathrm{g}, \mathrm{V})$ has ∞-Ricci curvature bounded below by K.

(iii) For all $U \in \mathrm{DC}_\infty, U_v$ is weakly $\lambda(U)$-displacement convex on $\mathrm{P}_2(M)$.

(iv) For all $U \in \mathrm{DC}_\infty, U_v$ is weakly $\lambda(U)$-a.c. displacement convex on $\mathrm{P}_2^{ac}(M)$

(v) $H_{\infty,\mathrm{V}}$ is weakly K -a.c. displacement convex on $\mathrm{P}_2^{ac}(M)$.

For both parts (a) and (b), the nontrivial implications are $(i) \Rightarrow (ii)$ and $(v) \Rightarrow (i)$. The proof that $(i) \Rightarrow (ii)$ will be along the lines of [159], with some differences. One ingredient the following lemma.

Proof : part (a). To show $(i) \Rightarrow (ii)$, suppose that $Ric_N \ge 0$. By the definition of Ric_N, we must have $n < N$, or $n = N$ and ϕ is constant. Suppose first that $n < N$. We can write

$$Ric_N = Ric - (N-n)e^{\frac{\phi}{N-n}} Hess\left(e^{-\frac{\phi}{N-n}}\right). \qquad (156)$$

Given $\mu_0, \mu_1 \in \mathrm{P}_2^{ac}(M)$,, let $\{\mu_t\}_{t\in[0,1]}$ be the unique Wasserstein geodesic from μ_0 to μ_1. From Proposition(4.1.29), in order to prove (ii) it suffices to show that for all such μ_0 and μ_1, and all $U \in \mathrm{DC}_N$, the inequality (37) in Section (4.1) is satisfied with $\lambda = 0$.

We recall facts from Section (4.1) about optimal transport on Riemannian manifolds. In particular, μ_t is absolutely continuous with respect to $\mathrm{dvol_M}$ for all t, and takes the form $(F_t)_* \mu_0$, where $F_t(y) = \exp_y(-t\nabla\phi(y))$ for some $\frac{d^2}{2}$ concave function ϕ. Put $\eta_t = \frac{d\mu_t}{dvol_M}$. Using the nonsmooth change-of-variables formula proven in [31] (see also [234]), we can write

$$U_v(\mu_t) = \int_M U\left(e^{\phi(m)}\eta_t(m)\right) e^{-\phi(m)} dvol_M(m)$$

$$= \int_M U\left(e^{\phi(F_t(y))} \frac{\eta_0(y)}{\det(dF_t)(y)}\right) e^{-\phi(F_t(y))} \det(dF_t)(y)\, dvol_M(y). \tag{157}$$

Putting

$$c(y,t) = e^{-\frac{\phi(F_t(y))}{N}} \det^{\frac{1}{N}}(dF_t)(y), \tag{158}$$

we can write

$$U_V(\mu_t) = \int_M c(y,t)^N U\left(\eta_0(y) c(y,t)^{-N}\right) dvol_M(y). \tag{159}$$

Suppose that we can show that $c(y,t)$ is concave in t for almost all $y \in M$. Then for $y \in \mathrm{supp}(\mu_0)$, as the map

$$\lambda \to \eta_0^{-1}(y) \lambda^N U\left(\eta_0(y)\lambda^{-N}\right) \tag{160}$$

is nonincreasing and convex, and the composition of a nonincreasing convex function with a concave function is convex, it follows that the integrand of (159) is convex in t. Hence $U_V(\mu_t)$ will be convex in t.

To show that $c(y,t)$ is concave in t fix y. Put

$$c_1(t) = e^{-\frac{\phi(F_t(y))}{N-n}} \tag{161}$$

and

$$c_2(t) = \det^{\frac{1}{n}}(dF_t)(y), \tag{162}$$

So $c(y,t)=c_1(t)^{\frac{N-n}{N}}c_2(t)^{\frac{n}{N}}$. We have

$$NC^{-1}\frac{d^2c}{dt^2}=(N-n)c_1^{-1}\frac{d^2c_1}{dt^2}+nc_2^{-1}\frac{d^2c_2}{dt^2}-\frac{n(N-n)}{N}\left(c_1^{-1}\frac{dc_1}{dt}-C_2^{-1}\frac{dc_2}{dt}\right)^2$$

$$\leq(Ric-Ric_N)\left(F_t'(y),F_t'(y)\right)+nc_2^{-1}\frac{d^2c_2}{dt^2}. \qquad (163)$$

We may assume that the function ϕ has a Hessian at y [31], and that dF_t is well-defined and nonsingular at y for all $t\in[0,1)$ [31] Then Lemma (4.2.25), shows that

$$nc_2^{-1}\frac{d^2c_2}{dt^2}\leq -Ric\left(F_t'(y),F_t'(y)\right). \qquad (164)$$

So $Nc^{-1}(t)C''c(t)\leq -Ric_N\left(F_t'(y),F_t'(y)\right)\leq 0$. This shows that (M,g,V) is weakly displacement convex for the family DC_N.

The proof in the case $N=n$ follows the same lines, replacing c_1 by 1 and c_2 by c.

We now prove the implication $(v)\Rightarrow(i)$. Putting $U=U_N$ in (77), we obtain

$$H_{N,v}(\mu_t)=N-N\int_M c(y,t)\eta_0(y)^{1-\frac{1}{N}}\,dvol_M(y). \qquad (165)$$

Suppose first that $n<N$ and $H_{N,\mathrm{V}}$ is weakly a.c. displacement convex. Given $m\in M$ and $\mathrm{V}\in T_mM$, we want to show that $Ric_N(\mathrm{V},\mathrm{V})\geq 0$. Choose a smooth function ϕ, defined in a neighborhood of m, so that $v=-(\nabla\phi)(m)$, Hess $(\phi)(m)$ is proportionate to $g(m)$ and

$$\frac{1}{N-n}\mathrm{V}\phi=\frac{1}{n}(\Delta\phi)(m). \qquad (166)$$

Consider the geodesic segment $t\rightarrow \exp_m(t\mathrm{V})$. Then

$$c_1^{-1}(0)c_1'(0)=-\frac{1}{N-n}\mathrm{V}\phi \qquad (167)$$

and

$$c_2^{-1}(0)c_1''(0)=\frac{1}{2n}Tr\left(W'(0)W^{-1}(0)\right)=\frac{1}{n}Tr\left(\mathbf{R}(0)\right) \tag{168}$$

$$=-\frac{1}{n}Tr\left(Hess(\phi)(m)\right)=-\frac{1}{n}(\Delta\phi)(m).$$

Hence by construction, $c_1^{-1}(0)c_1'(0)=c_2^{-1}(0)c_2'(0)$. From (164), it follows that

$$NC^{-1}(0)c''(0)=\left(Ric-Ric_N\right)(V,V)+nc_2^{-1}(0)c_2''(0). \tag{169}$$

As $R(0)$ is a multiple of the identity, (154) now implies that

$$Nc^{-1}(0)c''(0)=-Ric_N(V,V). \tag{170}$$

For small numbers $\varepsilon_1,\varepsilon_2>0$ consider a smooth probability measure μ_0 with support in an ε_1 ball around m. Put $\mu_1=\left(F_{\varepsilon_2}\right)_*\mu_0$ where F_t is defined by $F_t(y)=\exp_y\left(-t\nabla\phi(y)\right)$.. If ن$_2$ is small enough then $\varepsilon_2\phi$ is $\frac{d^2}{2}$-concave. As μ_0 is absolutely continuous, it follows that F_{ε_2} is the unique optimal transport between μ_0 and $\left(F_{\varepsilon_2}\right)_*\mu_0$. As a consequence, $\mu_t\equiv\left(F_{t_{ن2}}\right)_*\mu_0$ is the unique Wasserstein geodesic from μ_0 to μ_1. Taking $\varepsilon_t\to 0$ and then $\varepsilon_2\to 0$, if $H_{N,V}$ is to satisfy (36) in Section (4.1) for all such μ_0 then we must have $c''(0)\le 0$. Hence $Ric_N(V,V)\ge 0$. Since V was arbitrary, this shows that $Ric_N\ge 0$.

Now suppose that $N=n$ and $H_{N,V}$ is weakly a.c. displacement convex. Given

$m\in M$ and $V\in T_mM$, we want to show that $V\phi=0$ and $Ric(V,V)\ge 0$.. Choose a smooth function ϕ, defined in a neighborhood of m, so that $V=-(\nabla\phi)(m)$, and $\text{Hess}(\phi)(m)$ is proportionate to $g(m)$. We must again have $c''(0)\le 0$, where now $c(t)=e^{-\frac{\phi(F_t(y))}{n}}\det^{\frac{1}{n}}(dF_t)(y)$. By direct computation,

$$\frac{c''(0)}{c(0)}=-\frac{1}{n}\left(Ric+Hess(\phi)\right)(V,V)+\frac{(V\phi)^2}{n^2}+\frac{2(V\phi)(\Delta\phi)(m)}{n^2} \tag{171}$$

If $V\phi\neq 0$ then we can make $c''(0)>0$ by an appropriate choice of $\Delta\phi$. Hence ϕ must be constant and then we must have $Ric(V,V)\ge 0$.

Finally , if $N<n$

$$N\frac{c''(0)}{c(0)} = -\left(Ric + Hess(\phi)\right)(\mathrm{V},\mathrm{V}) + \frac{(\mathrm{V}\phi)^2}{N-n} - \frac{n(N-n)}{N}\left(-\frac{\mathrm{V}\phi}{N-n} + \frac{(\Delta\phi)(m)}{n}\right)^2. \quad (172)$$

One can always choose $(\Delta\phi)(m)$ to make $c''(0)$ positive, so $H_{N,\mathrm{V}}$ cannot be weakly a.c.

displacement convex.
□

part (b). We first show $(i) \Rightarrow (ii)$.suppose that $Ric_\infty \geq Kg$.

Given $\mu_0, \mu_1 \in \mathrm{P}_2^{ac}(M)$, we again use (157), with $U \in \mathrm{DC}_\infty$. Putting

$$c(y,t) = -\phi(F_t(y)) + \log\det(dF_t)(y), \quad (173)$$

we have

$$U_v(\mu_t) = \int_M e^{C(y,t)} U\left(\eta_0(y) e^{-C(y,t)}\right) dvol_M(y). \quad (174)$$

As in the proof of (a), the condition $Ric_\infty \geq Kg$ implies that

$$\frac{d^2c}{dt^2} \leq -K\left|F_t'(y)\right|^2 = -K\left|\nabla\phi\right|^2(y), \quad (175)$$

where the last equality comes from the constant speed of the geodesic $t \to F_t(y)$. By assumption, the map

$$\lambda \to \eta_0^{-1}(y) e^{\lambda} U\left(\eta_0(y) e^{-\lambda}\right) \quad (176)$$

is nonincreasing and convex in λ, with derivative $\frac{p\left(\eta_0(y)e^{-\lambda}\right)}{\eta_0(y)e^{-\lambda}}$._It follows that the composition

$$\lambda \to \eta_0^{-1}(y) e^{C(y,t)} U\left(\eta_0(y) e^{C(y,t)}\right) \quad (177)$$

is $\lambda(U)|\nabla\phi|^2(y)$-convex in t. Then

$$e^{C(y,t)} U\left(\eta_0(y) e^{-C(y,t)}\right) \leq t e^{C(y,1)} U\left(\eta_0(y) e^{-C(y,1)}\right)$$

$$+(1-t) e^{C(y,0)} U\left(\eta_0(y) e^{-C(y,0)}\right) - \frac{1}{2}\lambda(U)|\nabla\phi|^2(y)\eta_0(y)t(1-t). \quad (178)$$

Integrating with respect to $\mathrm{dvol_M}(y)$ and using the fact that

$$W_2(\mu_0,\mu_1)^2 = \int_M |\nabla\phi|^2(y)\eta_0(y)\,dvol_M(y) \tag{179}$$

shows that (37) in section (4.1)is satisfied with $\lambda = \lambda(U)$. The implication $(i) \Rightarrow (ii)$ now follows from Proposition(4.1.29).

The proof that $(v) \Rightarrow (i)$ is similar to the proof in part (a).

The case $N = 1$ is slightly different because $H_{1,\mathrm{V}}$ is not defined. However, the rest of

Theorem (4.2.26)a carries through.

Theorem (4.2.27)[121]: (a)The following are equivalent:

(i) $Ric_1 \geq 0$..

(ii) The measured length space (M, g, V) has nonnegative 1-Ricci curvature.

(iii) For all $U \in DC_1, U_v$ is weakly displacement convex on $P_2(M)$.

(iv) For all $U \in DC_1, U_v$ is weakly a.c. displacement convex on $P_2^{ac}(M)$.

Corollary (4.2.28)[121]: Let (B, g_B) be a smooth compact connected Riemannian manifold, equipped with the Riemannian density $\mathrm{dvol_B}$, and let ϕ be a C^2-regular function on B which is normalized by an additive constant so that $e^{-\phi}$ $\mathrm{dvol_B}$ is a probability measure on B. We have the following implications:

(i) If $(B, gB, e^{-\Psi} dvol_B)$ is a measured Gromov-Hausdorff limit of Riemannian manifolds with nonnegative Ricci curvature and dimension at most N then $Ric_N(B) \geq 0$.

(ii) If $(B, gB, e^{-\Psi} dvol_B)$is a measured Gromov-Hausdorff limit of Riemannian manifolds with Ricci curvature bounded below by $K \in \mathbf{R}$ then $Ric_\infty(B) \geq KgB$..

(iii) As a partial converse, if $(B, gB, e^{-\Psi} dvol_B)$ has $Ric_N(B) \geq 0$ with $N \geq \dim(B) + 2$

an integer then $(B, gB, e^{-\Psi} dvol_B)$ is a measured Gromov-Hausdorff limit of Riemannian manifolds with nonnegative Ricci curvature and dimension at most N.

(iv) If $(B, gB, e^{-\Psi} dvol_B)$ has $Ric_\infty(B) \geq KgB$. then $(B, gB, e^{-\Psi} dvol_B)$ is a measured

Gromov-Hausdorff limit of Riemannian manifolds with $Ric(M_i) \geq \left(K - \frac{1}{i}\right) g_{M_i}$

Corollary (4.2.29)[121]: (a) Suppose that (X, d) is a Gromov-Hausdorff limit of

n-dimensional Riemannian manifolds with nonnegative Ricci curvature. If (X, d) has Hausdorff dimension n, and ν_H is its normalized n-dimensional Hausdorff measure, then (X, d, ν_H) has nonnegative n- Ricci curvature.

(b) If in addition (X, d) happens to be a smooth n-dimensional Riemannian manifold B, gB then $Ric(B) \geq 0$.

proof: (a)If $\{M_i\}_{i=1}^\infty$ is a sequence of n-dimensional Riemannian manifolds with nonnegative Ricci curvature and $\{f_i\}_{i=1}^\infty$ is a sequence of ε_1- approximations $f_i : M_i \to X$, with $\lim_{i\to\infty} \varepsilon_i = 0$, then $\lim_{i\to\infty} (f_i)_* dvol_{M_i} = \nu_H$ in the weak-* topology [108]. (This also shows that the n-dimensional Hausdorff measure on X can be normalized to be a probability measure.) Then part a. follows from Theorems (4.2.8) and Theorems (4.2.26)

(b) If $(X, d) = (B, g_B)$ then $\nu_H = \frac{dvol_B}{vol(B)}$ and the claim follows from Theorem (4.2.26): along with the definition of Ric_n.

(X, d) has nonnegative Alexandrov curvature then (X, d, ν_H) has nonnegative n-Ricci curvature. For $n > 1$, if (X, d) has Alexandrov curvature bounded below by $\frac{k}{n-1}$ then (X, d, ν_H) has ∞- Ricci curvature bounded below by *K*.

As mentioned above, in the collapsing case the lower bound in the conclusion of Corollary (4.2.28)[(i) (or Corollary (4.2.28) (ii)) would generally fail if we replaced Ric_N (or Ric_∞) by Ric. However, one does obtain a lower bound on the average scalar curvature of *B*.

Corollary (4.2.30)[121]: If $(B, gB, e^{-\Psi} dvol_B)$ is a smooth n-dimensional measured Gromov -Hausdorff limit of Riemannian manifolds (of arbitrary dimension), each with Ricci curvature bounded below by $K \in \mathbf{R}$, then the scalar curvature S of (B, gB) satisfies

$$\frac{\int_B S\, dvol_B}{vol(B)} \geq nK. \tag{180}$$

Corollary(4.2.31)[121]: Let M be a compact connected Riemannian manifold. Let G be a compact Lie group that acts isometrically on M, preserving a function $\Psi \in C^\infty(M)$ that satisfies $\int_M e^{-\Psi} dvol_M = 1$. Let $p : M \to M/G$ be the quotient map.

a. For $N \in [1,\infty)$, if $(M, e^{-\Psi} dvol_M)$ has $Ric_N \geq 0$ then $(M/G, d_{M/G,} p_*(e^{-\Psi} dvol_M))$ has nonnegative N-Ricci curvature.

b. If $(M, e^{-\Psi} dvol_M)$ has $Ric_\infty \geq Kg_M$ then $M/G, d_{M/G,} p_*(e^{-\Psi} dvol_M)$ has ∞-Ricci curvature bounded below by k.

Corollary (4.2.32)[121]: provides many examples of singular spaces with lower Ricci curvature bounds. Of course, the main case is when Ψ is constant.

We conclude this section by giving a "synthetic" proof of a part of the Ricci O'Neill theorem of [119].

Corollary(4.2.33)[121]: Let $p : M \to B$ be a Riemannian submersion of compact connected manifolds, with fibers Z_b.Choose $N \geq \dim(M)$ and $\Psi_M \in C^\infty(M)$ with $\int_M e^{-\Psi_M} dvol_M = 1$; if $N = \dim(M)$ then we assume that Ψ_M is constant. Define $\Psi_B \in C^\infty(B)$ by $p_*(e^{-\Psi_M} dvol_M) = e^{-\Psi_B} dvol_B$. Suppose that the fiber parallel transport of the Riemannian submersion preserves the fiberwise measures $e^{-\Psi_M}\big|_z dvol_z$ up to multiplicative constants. (That is, if $\gamma : [0,1] \to B$ is a smooth path in B, let $p_\gamma : Z_{\gamma(0)} \to Z_{\gamma(1)}$ denote the fiber transport diffeomorphism. Then we assume that there is a constant $c_\gamma > 0$ so that

$$p_\gamma^*\left(e^{-\psi_M}\Big|_{Z_{\gamma(1)}} dvol_{Z_{\gamma(1)}}\right) = c_\gamma e^{-\psi_M}\Big|_{Z_{\gamma(1)}} dvol_{Z_{\gamma(0)}}.) \tag{181}$$

With these assumptions,

a. If $Ric_N(M) \geq 0$ then $Ric_N(B) \geq 0$.

b. For any $K \in \mathbf{R}$, if $Ric_\infty(M)dvol_{Z_{\gamma(1)}} \geq Kg_M$ then $Ric_\infty(B) \geq Kg_M$.

Proof: Put $V_M = e^{-\Psi_M} dvol_M$ and $V_B = e^{-\Psi_B} dvol_B$ We can decompose V_M with respect to p as $\sigma(b)V_B(b)$, with $\sigma(b) \in P_2^{ac}(Z_b)$ From the assumptions, the family $\{\sigma(b)\}_{b\in B}$ of vertical densities is invariant under fiber parallel transport.

To prove part (a), let $\{\mu_t\}_{t\in[0,1]}$ be a Wasserstein geodesic in P_2^{ac}. Define $\{\mu_t'\}_{t\in[0,1]}$ in $P_2^{ac}(M)$ by $\mu_t' \equiv \sigma(b)\mu_t(b)$.By construction, the corresponding densities satisfy $\rho_t' = p^*\rho_t$. Thus $H_{N,v_M}(\mu_t') = H_{N,v_B}(\mu_t)$. Furthermore, $\{\mu_t'\}_{t\in[0,1]}$ is a Wasserstein geodesic; if $(F_t)_{t\in[0,1]}$ is an optimal Monge transport from μ_0 to μ_1 then its horizontal lift is an optimal Monge transport from μ_0' to μ_1', with generating function $\phi_M = p^*\phi_B$ From Theorem(4.2.26) (a) H_{N,v_M} is a.c. displacement convex on $P_2^{ac}(M)$. In particular, (36) in section (4.1) is satisfied along $\{\mu_t'\}_{t\in[0,1]}$ with $U_v = H_{N,v_M}$ and $\lambda = 0$. Then the same equation is satisfied along $\{\mu_t\}_{t\in[0,1]}$ with H_{N,v_B} and $\lambda = 0$. Thus $\{\mu_t\}_{t\in[0,1]}$ is a.c.

displacement convex on $P_2^{ac}(B)$.. Theorem (4.2.26) (a) now implies that $Ric_N(B) \geq 0$.

The prove of part (b) is similar.

Section (4.3): Mass transportation and rough curvature bounds

We develop a notion of rough curvature bounds for discrete spaces, based on the concept of optimal mass transportation. These rough curvature bounds will depend on a real parameter $h > 0$, which should be considered as a natural length scale of the underlying discrete space or as the scale on which we have to look at the space. For a metric graph, for instance, this parameter equals the maximal length of its edges (times some constant).

. For instance, instead of midpoints of a given pair of points x_0, x_1 we look at $h-$midpoints which are points y with $d(x_0,y) \leq \frac{1}{2}d(x_0,x_1) + h$ and $d(x_1,y) \leq \frac{1}{2}d(x_0,x_1) + h$.

Given any metric space (M,d,m) with curvature $\geq K$ and any $h>0$ we define standard discretizations (M_h,d,m_h) of (M,d,m) with $D^2((M_h,d,m_h),(M,d,m))\to 0$ as $h\to 0$ and with $h-Curv(M,d,m)\geq K_h$.

Throughout this section, a metric measure space will always be a triple (M,d,m) where (M,d) is a complete separable metric space and m is a measure on M (equipped with its Borel $\sigma-$algebra $\mathrm{B}(M)$) which is locally finite in the sense that $m(B_r(x))<\infty$ for all $x\in M$ and all sufficiently small $r>0$. We say that the metric measure space (M,d,m) is normalized if $m(M)=1$. Two metric measure spaces (M,d,m) and (M',d',m') are called isomorphic if and only if there exists an isometry $\psi: M_0\to M_0'$ between the supports $M_0:=\mathrm{supp}[m]\subset M$ and $M_0':=\mathrm{supp}[m']\subset M'$, such that $\psi_*m=m'$. The diameter of a metric measure space (M,d,m) will be the diameter of the metric space $(\mathrm{supp}[\mathrm{m}],\mathrm{d})$.

We shall use the notion of L $_2$-transportation distance D for two metric measure spaces (M,d,m) and (M',d',m'), as defined in [141]:

$$D(M,d,m),(M',d',m')=\inf\left(\int_{M\bigcup M'}\hat{d}^2(x,y)\,dq(x,y)\right)^{1/2}$$

where $\hat{\mathrm{d}}$ ranges over all couplings of d and d' and q ranges over all couplings of m and m'. Here a measure q on the product space $M\times M'$ is a coupling of m and m' if $q(A\times M')=m(A)$ and $q(M\times A')=m'(A')$ for all measurable $A\subset M, A'\subset M'$; a pseudo-metric $\hat{\mathrm{d}}$ on the disjoint union $M\bigcup M'$ is a coupling of d and d' if $\hat{d}(x,y)=d(x.y)$^)and $\hat{d}(x',y')=d'(x',y')$ for all $x,y\in\mathrm{supp}[m]\subset M$ and all $x',y'\in\mathrm{supp}[m']\subset M'$.

The L_2-transportation distance D defines a complete and separable length metric on the family of all isomorphism classes of normalized metric measure spaces (M,d,m) for which $\int_M d^2(z,x)\,dm(x)<\infty$ some (hence all) $z\in M$. The notion of $D-$convergence is closely related to the one of measured Gromov–Hausdorff convergence introduced in [128]. Recall that a sequence of compact normalized metric measure spaces

$\{(M_n,d_n,m_n)\}_{n\in N}$.converges in the sense of *measured* Gromov–Hausdorff convergence (briefly, mGH -converges) to a compact normalized metric measure space (M,d,m) iff there exist a sequence of numbers $\varepsilon\,]\;\;0$ and a sequence of measurable maps $f_n : M_n \to M$ such that for all $x, y \in M_n, \left| d\left(f_n(x), f_n(y)\right) - d_n(x, y) \right| \le \varepsilon_n$, for any $x \in M$ there exists $y \in M_n$ with

$d\left(f_n(y), x\right) \le \varepsilon_n$ and such that $(f_n)_* m_n \to m$. weakly on M for $n \to \infty$. According to Lemma (1.2.)[141]: any m GH-convergent sequence of normalized metric measure spaces is also D-convergent; for any sequence of normalized compact metric measure spaces with full supports and with uniform bounds for the doubling constants and for the diameters the notion of $m\,GH-\text{convergence}$ is equivalent to the one of D -converg-ence. It is easy to see that $D\left((M,d,m),(M',d',m')\right) = \inf \hat{W}\left(\phi_* m, \phi'_* m'\right)$ where the inf is taken over all metric spaces $\left(\hat{M}, \hat{d}\right)$ with isometric embeddings $\phi : M_0 \to \hat{M}, \phi' : M'_0 \to \hat{M}$

of the supports M_0 and M'_0 of m and m', respectively, and where $d\hat{W}$ denotes the

L_2 -Wasserstein distance derived from the metric $\hat{d}$. Recall that for any metric space (M,d) the L_2-Wasserstein distance between two measures μ and V on M is defined as

$$W(\mu, V) = \inf\left\{\left(\int_{M\times M} d^2(x,d)\, dq(x,y)\right)^{1/2} : q \text{ isa coupling of } \mu \text{ and } V\right\},$$

with the convention $\inf \theta = \infty$.. For further details about the Wasserstein distance see the monograph [23]. We denote by $P_2(M,d)$ the space of all probability measures V which have finite second moments $\int_M d^2(o,x)\, dv(x) < \infty$ for some (hence all) $o \in M$. For a given metric measure space (M,d,m) we put $P_2(M,d,m)$ the space of all probability measures $v \in P_2(M,d)$ which are absolutely continuous w. r. t. $V\,m$. If $V = \rho.m \in P_2(M.d.m)$ we consider the *relative entropy* of V with respect to m defined by $H(V|m) := \lim_{\varepsilon]\ 0} \int_{\{\rho > \varepsilon\}} \rho \log \rho\, dm$. We denote by $P_2^*(M,d,m)$ the subspace of measures $V \in P_2(M,d,m)$ of finite entropy $H(V|m) < \infty$.

We recall here the definitions of the lower curvature bounds for metric measure spaces introduced in [141]:

A metric measure space (M,d,m) has *curvature* $\ge K$ for some number $K \in R$ if and only if the relative entropy $H(.|m)$ is weakly K -convex on

$P_2^*(M,d,m)$ in the sense that for each pair $\nu_0,\nu_1 \in P_2^*(M,d.m)$ there exists a geodesic $\Gamma:[0,1]\to P_2^*(M,d,m)$ connecting v_0 and ν_1 with

$$H(\Gamma(t)|m) \le (1-t)H(\Gamma(0)|m) + tH(\Gamma(1)|m) - \frac{K}{2}t(1-t)W^2(\Gamma(0),\Gamma(1)) \quad (182)$$

for all $t\in[0,1]$.

The metric measure space (M,d,m) has *curvature* $\ge K$ in the lax sense if and only if for each $\varepsilon>0$ and for each pair $v_0,v_1 \in P_2^*(M,d,m)$ here exists an $\varepsilon-$midpoint $\eta\in P_2^*(M,d,m)$ of ν_0 and ν_1 with

$$H(\eta|m) \le \frac{1}{2}H(\nu_0|m) + \frac{1}{2}H(\nu_1|m) - \frac{K}{8}W^2(\nu_0,\nu_1) + \varepsilon \quad (183)$$

Briefly, we shall write $Curv(M,d,m)\ge K$, respectively $Curv_{lax}(M,d,m)\ge K$.

Recall that in a given metric space $(M,d\ )$ a point y is an ε-midpoint of x_0 and x_1 if $d(x,y)\le\frac{1}{2}d(x_0,x_1)+\varepsilon$ for each $i=0,1$. We call y *midpoint* of x_0 and x_1 if $d(x,y)\le\frac{1}{2}d(x_0,x_1)$ for $i=0,1$.

In order to adapt the notion of curvature bound to other spaces then geodesic without branching we shall refer in this section to a larger class of metric spaces:

Definition (4.3.1)[7]: Let $h>0$ be given. We say that a metric space $(M,d\ )$ is $\text{h}-\text{rough}$ geodesic iff for each pair of points $x_0,x_1\in M$ and each $t\in(0,1)$ there exists a point $x_t\in M$ satisfying

$$d(x_0,x_1)\le td(x_0,x_1)+h, d(x_t,x_1)\le(1-t)d(x_0,x_1)+h \quad (184)$$

The point x_t will be referred to as the $h-\text{rought}-\text{approximate}$ point between $x_0, and\ x_1$ The $h-\text{rought}\frac{1}{2}-\text{approximate}$ point is actually the $h-\text{midpoint}$ of $x_0, and\ x_1$

Example (4.3.2)[7]: Any nonempty set X with the discrete metric $d(x,y)=0$ for $x=y$ and 1 for $x\ne y$ h-rough geodesic for any $h\ge\frac{1}{2}$.In this case, any point is an $h-\text{midpoint}$ of any pair is distinct points.

If $\varepsilon>0$ then the space (R^n,d) with the metric $d(x,y)=|x-y|\wedge\varepsilon$.is $\text{h}-\text{rough}$ geodesic for $h\ge\varepsilon/2$ (here $|\cdot|$is the Euclidian metric).

(iii) For $\varepsilon > 0$ the space (R^n, d) with the metric $d(x,y) = \sqrt{\varepsilon|x-y| + |x-y|^2}$ is $h-$rough geodesic for each $h \geq \varepsilon / 4$.

The above examples are somewhat pathological. We actually have in mind the more friendly examples of discrete spaces and some geodesic spaces with branch points, e.g. graphs, that do not have curvature bounds as defined in [141].

For a discrete $h-$rough geodesic metric space (M,d) one should think of h as a discretization size or "resolution" of M .In an $h-$geodesic space a pair of points x and y is not necessarily connected by a geodesic but by a chain of points $x = x_0, x_1, ..., x_n = y$ having intermediate distance less then $\frac{h}{2}$.In the sequel we will use two types of perturbations of the Wasserstein distance, defined as follows:

Definition (4.3.3)[7]: Let (M,d) be a metric space. For each $h > 0$ and any pair of measures $\nu_0, \nu_1 \in P_2(M,d)$ put

$$W^{\pm h}(\nu_0,\nu_1) := \inf\left\{\left(\int\left[\left(d(x_0,x_1) \mp h\right)_+\right]^2 dq(x_0,x_1)\right)^{1/2} : q \text{ coupling of } \nu_0 \text{ and } \nu_1\right\}, \quad (185)$$

where (.) + denotes the positive part.

The two perturbations W^{+h} and W^{-h} are related to the Wasserstein distance W in the following way

Lemma(4.3.4)[7]: For any $h > 0$ we have

$(i)\ W^{+h} \leq W \leq W^{+h} + h;\ (ii)\ W \leq W^{-h} \leq W + h$.

Proof: (i) Let ν_0 and ν_1 be two probabilities in (M,d) and consider q an optimal coupling and $q+h$ $a+h$-optimal coupling of them. Then

$$W^{+h}(\nu_0,\nu_1) = \left(\int\left[\left(d(x_0,x_1) - h\right)_+\right]^2 dq + h(x_0,x_1)\right)^{1/2} \leq \left(\int\left[\left(d(x_0,x_1) - h\right)_+\right]^2 dq + h(x_0,x_1)\right)^{1/2}$$

$$\leq \left(\int d(x_0,x_1)^2\, dq(x_0,x_1)\right)^{1/2} = W(\nu_0,\nu_1)$$

and

$$W(\nu_0,\nu_1) = \left(\int d(x_0,x_1)^2\, dq(x_0,x_1)\right)^{1/2} \leq \left(\int d(x_0,x_1)^2\, dq + h(x_0,x_1)\right)^{1/2}$$

$$\leq\left(\int\left[\left(d\left(x_0,x_1\right)-h\right)_+ + h\right]^2 dq + h\left(x_0,x_1\right)\right)^{1/2} \leq W^{+h}\left(V_0,V_1\right)+h$$

(ii) Similar to (i).

With an elementary proof we have also a monotonicity property of $W^{\pm h}$ in h :

Lemma(4.3.5)[7]: Let $0<h_1<h_2$ be arbitrarily given. Then for each pair of probabilities v_0 and v_1

(i) $W^{-h_1}\left(V_0,V_1\right)<W^{-h_2}\left(V_0,V_1\right)$;

(ii) $W^{+h_1}\left(V_0,V_1\right)\geq W^{+h_2}\left(V_0,V_1\right)$ and the inequality is strict if and only if $W^{+h_1}\left(V_0,V_1\right)>0.$

We introduce now the notion of rough lower curvature bound:

Definition(4.3.6)[7]: We say that a metric measure space (M,d,m) has h-*rough curvature* $\geq K$ for some numbers $h>0$ and $K\in R$ iff for each pair $V_0,V_1\in P_2^*(M,d.m)$ and for any $t\in[0,1]$ here exists an h -rough t -approximate point $\eta_t\in P_2^*(M,d,m)$ between v_0 and v_1 satisfying

$$H\left(\eta_t|m\right)\leq(1-t)H\left(V_0|m\right)+tH\left(V_1|m\right)-\frac{K}{2}t(1-t)W^{\pm h}\left(V_0,V_1\right)^2 \tag{186}$$

where the sign in $W^{\pm h}\left(V_0,V_1\right)$ is chosen ' +' if $K>0$ and ' -' if $K<0$. Briefly, we write in this case $h-curv(M,d,m)\geq K$.

Corollary (4.3.8)[274]: If any $h_n>0$ we have

$(i)\ W^{+h_n}\leq W\leq W^{+h_n}+h_n\ (ii)\ W\leq W^{-h_n}\leq W+h_n$.

Proof: (i) Let V_n and $V_{n+1}, n\geq 0$ be two probabilities in (M,d). Now consider q an optimal coupling and $q+h_n, a+h_n$ -optimal coupling of them. Then

$$W^{+h}\left(V_n,V_{n+1}\right)=\left(\int\left[\left(d\left(x_n,x_{n+1}\right)-h_n\right)_+\right]^2 dq+h_n\left(x_n,x_{n+1}\right)\right)^{1/2}\leq\left(\int\left[\left(d\left(x_n,x_{n+1}\right)-h_n\right)_+\right]^2 dq+h_n\left(x_n,x_{n+1}\right)\right)^{1/2}$$

$$\leq\left(\int d\left(x_n,x_{n+1}\right)^2 dq\left(x_n,x_{n+1}\right)\right)^{1/2}=W\left(V_n,V_{n+1}\right)$$

and

$$W\left(V_n,V_{n+1}\right)=\left(\int d\left(x_n,x_{n+1}\right)^2 dq\left(x_n,x_{n+1}\right)\right)^{1/2}\leq\left(\int d\left(x_n,x_{n+11}\right)^2 dq+h\left(x_n,x_{n+1}\right)\right)^{1/2}$$

$$\leq\left(\int\left[\left(dx_{n},x_{n+1}\right)-h_{n}\right)_{+}+h_{n}\Big]^{2}dq+h_{n}\left(x_{n},x_{n+1}\right)\right)^{1/2}\leq W^{+h}\left(V_{n},V_{n+1}\right)+h_{n}$$

(ii) Similar to be find as (i).

Theorem (4.3.9)[7]: Let (M,d,m) be a normalized metric measure space and $\{(M_h,d_h,m_h)\}_{h>0}$ a family of normalized metric measure spaces with uniformly bounded diameter and with $h-curv(M_h,h_h,m_h)\geq K_h$ for $K_h\to K$ as $h\to 0$ if

$$(M_h,d_h,m_h)\overset{D}{\to}(M,d,m)$$

as $h\to 0$ then

$$curv_{lax}(M,d,m)\geq K.$$

If in addition M is compact then

$$curv(M,d,m)\geq K.$$

Proof: Let $\{(M_h,d_h,m_h)\}_{h>0}$ be a family of normalized discrete metric measure spaces. Assume that $(M_h,d_h,m_h)\xrightarrow{D}(M,d,m)$ as $h\to 0$ and $\sup_{h\to 0} diam(M_h,d_h,m_h), diam(M,d,m)\leq\Delta$ for some $.\Delta\in R$ .Now let $\varepsilon>0$ and $V_0=\rho_o m, V_1=\rho_1 m\in P_2^*(M,d,m)$ be given. Choose **R** with

$$\sup_{i-0,1}H(v_i|m)+\frac{|K|}{8}\Delta^2+\frac{\varepsilon}{8}\left[\Delta^2+3|K|(2\Delta+3\varepsilon)\right]\leq R. \qquad (187)$$

We have to deduce the existence of an ε-midpoint η which satisfies inequality (2). Choose $0<h<\varepsilon$ with $|K_h-K|<\varepsilon$ and

$$D(M_h,d_h.m_h),(M,d.m)\leq\exp\left(\frac{2+4\Delta^2R}{\varepsilon^2}\right) \qquad (188)$$

Like in [141],one can define the canonical maps $Q_h':P_2(M,d,m)\to P_2(M_h,d_h,m_h)$ and $Q_h:P_2(M_h,d_h,m_h)\to P_2(M,d,m)$ as follows.

We consider q_h a coupling of m and m_h and d_h a coupling of d and d_h such that

$$\int\hat{d}_h^2(x,y)\,dqh(x,y)\leq 2D^2\left((M,d.m),(M_h,d_h.m_h)\right)$$

Let Q'_h and Q_h be the disintegrations of q_h w. r. t. m_h and m , resp., that is $sqh(x,y) = Q'_h(y,dx)dm_h(y) = Q_h(x,dy)dm(x)$ and let $\hat{\Delta}$ denote the $\mathrm{m-essential}$ supremum of the map

$$x \mapsto \left[\int_{M_h} \hat{d}_h^2(x,y)Q_h(x,dy)\right]^{1/2}$$

In our case $\hat{\Delta} \le 2\Delta.$

For $\mathrm{V} = \rho m \in \mathrm{P}_2(M,d,m)$ define $Q'_h(\mathrm{V}) \in \mathrm{P}_2(M_h,d_h,m_h)$ by $Q'_h(\mathrm{V}) := \rho_h m_h$ where

$\rho_h(y) := \int_M \rho(x) Q'_h(y.dx)$

The map Q_h is defined similarly. From [12] gives the following estimates:

$$H\left(Q'_h(\mathrm{V})\middle|m_h\right) \le H\left(\mathrm{V}\middle|m\right) \text{ for all } \mathrm{V} = \rho m \tag{189}$$

and

$$\mathrm{W}^2\left(\mathrm{V}, Q'_h(\mathrm{V})\right) \le \frac{2 + \hat{\Delta}^2 H\left(\mathrm{V}\middle|m\right)}{-\log D(M,d,m),(M_h,d_h,m_h)} \tag{190}$$

provided $D(M,d,m),(M_h,d_h,m_h) < 1$. Analogous estimates hold for Q_h.

For our given $\mathrm{V}_0 = \rho o m, \mathrm{V}_1 = \rho_1 m \in \mathrm{P}_2^*(M,d,m)$ put

$\mathrm{V}_{i,h} := Q'_h(\mathrm{V}_i) = \rho_{i,h} m_h$

with $\rho_{i,h}(y) = \int \rho_i(x) Q'_h(y,dx)$ for $i = 0,1$ and let η_h be an $\mathrm{h-midpoint}$ of $\mathrm{V}_{0,h}$ and $\mathrm{V}_{1,h}$ such that

$$H\left(\eta_h\middle|m_h\right) \le \frac{1}{2} H\left(\mathrm{V}_{0,h}\middle|m_h\right) + \frac{1}{2} H\left(\mathrm{V}_{1,h}\middle|m_h\right) - \frac{K_h}{8} \mathrm{W}^{\delta_h h}\left(\mathrm{V}_{0,h}, \mathrm{V}_{1,h}\right)^2 \tag{191}$$

where δ_h is the sign of K_h

From (188)-(190) we conclude

$$\mathrm{W}^2\left(\mathrm{V}_0, \mathrm{V}_{0,h}\right) \le \frac{2 + \hat{\Delta}^2 . H\left(\mathrm{V}_0\middle|m\right)}{-\log D\left((M,d,m)(M_h,d_h,m_h)\right)} \le \frac{2 + \Delta^2 R}{-\log D\left((M,d,m),(M_h,d_h,m_h)\right)} \le \varepsilon^2$$

and similarly $W^2\left(V_1, V_{0,h}\right) \le \varepsilon^2$

If $K < 0$ we can suppose $K_h < 0$ too. From Lemma (4.3.5) (ii) we have

$$W^{-h}\left(V_{0,h}, V_{1,h}\right)^2 \le \left(W\left(V_{0,h}, V_{1,h}\right) + h\right)^2$$

$$\le \left(W\left(V_0, V_1\right) + 3\varepsilon\right)^2 \le W\left(V_0, V_1\right)^2 + 6\varepsilon\Delta + 9\varepsilon^2$$

because $W\left(V_0, V_{1,}\right) \le \Delta$.For $K < 0$ one can choose h small enough to ensure $K_h < 0$. Then Lemma (4.3.5)(i): implies.

$$W\left(V_0, V_1\right)^2 \le \left(W\left(V_{0,h}, V_{1,h}\right) + 2\varepsilon\right)^2 \le \left(W^{+h}\left(V_{0,h}, V_{1,h}\right) + 3\varepsilon\right)^2 \le W^{+h}\left(V_0, V_1\right)^2 + 6\varepsilon\Delta + 9\varepsilon^2$$

In both cases the estimates above combined with (189), (191) and the fact that we chose h i[th] $-K_h < \varepsilon - K$ will imply

$$H\left(\eta_h | m_h\right) \le \frac{1}{2} H\left(V_0 | m\right) + \frac{1}{2} H\left(V_1 | m\right) - \frac{K}{8} W^2\left(V_0, V_1\right) + \varepsilon' \qquad (192)$$

with $\varepsilon' = \varepsilon \dfrac{\left|\Delta^2 + 3\left|K\right|\left(2\Delta + 3\varepsilon\right)\right|}{8}$

The case $K = 0$ follows by the calculations above, depending on the sign of K_h.

Finally, put $\eta = Q_h\left(\eta_h\right)$.

Then again by (188), the estimates given in [141] for Q_h and by the previous estimate (192) for $H\left(\eta_h | m_h\right)$ we deduce

$$W^2\left(\eta_0, \eta\right) \le \frac{2 + \hat{\Delta}^2 . H\left(\eta_h | m_h\right)}{-\log D\left(\left(M, d, m\right), \left(M_h, d_h, m_h\right)\right)} \le \frac{2 + 4\Delta^2 R}{-\log\left(\left(M, d, m\right), \left(M_h, d_h, m_h\right)\right)} \le \varepsilon^2$$

For, i = 0,1we have $W\left(\eta, V_i\right) \le 2\varepsilon + +h \le \frac{1}{2} W\left(\eta_0, V_1\right) + 4\varepsilon$

Hence,

$$\sup_{i=0,1} W\left(\eta, V_i\right) \le \frac{1}{2} W\left(V_0, V_1\right) + 4\varepsilon$$

i.e. η is $a(4\varepsilon)-\text{midpoint}$ of V_0 and V_1. Furthermore, by (189)

$$H(\eta|m) \le H(\eta_h|m_h) \le \frac{1}{2}H(V_0|m) + \frac{1}{2}H(V_1|m) - \frac{K}{8}W^2(V_0,V_1) + \varepsilon'$$

with ε' as above. This proves that $Curv_{lax}(M,d,m) \ge K.$

Let (M,d,m) be a given metric measure space. For $h>0$ let be a discrete subset of M, say $M_h = \{x_n : n \in N\}$, with $M = \bigcup_{i=1}^{\infty} B_R(x_i)$, where $R = R(h) \to 0$ as $h \to 0$. If (M,d,m)

has finite diameter then M_h might consist of a finite number of points. Choose $A_i \subset B_R(x_i)$ mutually disjoint with $x_i \in A_i, i = 1,2,... and \bigcup_{i=1}^{\infty} A_i = M$ and (e.g. one could choose a Voronoi tessellation) and consider the measure M_h on given by $m_h(\{x_i\}) := m(A_i), i = 1,2,...$ We call (M_h,d_h,m_h) a discretization of (M,d,m).

Theorem (4.3.10)[7] (i)**:** If $m(M) < 0$ then $(M_h,d,m_h) \xrightarrow{D} (M,d,m)\ as\ h \to 0$

If $Curv_{lax}(M,d,m) \ge K.$ with $K \ne 0$ then for each $h > 0$ and for each discretization (M_h,d,m_h) with $R(h) < h/4$ we have $h - Curv(M_h,d,m_h) \ge K..$

If $Curv(M,d,m) \ge K.$ for some real number K then for each $h > 0$ and for each discretization (M_h,d,m_h) with $R(h) < h/4$ we have $h - Curv(M_h,d,m_h) \ge K.$h

Proof: (i) The measure $q = \sum_{i=1}^{\infty}(m(A_i)\delta_{x_i}) \times (1A_i m)$ is a coupling of m_h and m , so

$$D^2(M_h,d,m_h),(M,d,m) \le \int_{M_h \times M} d^2(x,y)dq(x,y) = \sum_{i=1}^{\infty} m(A_i)\int_{A_i} d^2(x,y)dm(y).$$

$$\le \sum_{i=1}^{\infty} m(A_i)R(h)^2 \le R(h)^2\left(\sum_{i=1}^{\infty} m(A_i)\right)^2$$

$$= R(h)^2 m(M)^2 \to 0 \text{ as } h \to 0.$$

(ii) Fix $h > 0$ and consider a discretization $(M_h,d,m_h),(M,d,m)$ with $R(h) < h/4$ Let $V_0^h, V_1^h \in P_2^*(M_h,d,m_h)$ be given; it is enough to make the proof for V_0^h, V_1^h with compact support. Suppose then $V_i^h = \left(\sum_{j=1}^{n} a_{i,j}^h 1_{\{x_j\}}\right) m_h, i = 1,2$

(some of the $a_{i.j}^{h}$ can be zero). We take also an arbitrary $t\in[0,1]$ Put $\nu_i:=\left(\sum_{j=1}^{n}a_{i,j}^{h}\right)m\in P_2^*(M,d,m)$ for $i=1,2$. Choose $\varepsilon>0$ such that

$$4R(h)+\varepsilon\leq h. \tag{193}$$

Since $Curv_{lax}(M,d,m)\geq K.$ for our given $t\in[0,1]$ there exists $\eta_t\in P_2^*(M,d,m)$ an ε-rough t -approximate point between ν_0 and ν_1 such that

$$H(\eta_t|m)\leq(1-t)H(\nu_0|m)+tH(\nu_1|m)-\frac{K}{2}t(1-t)W^2(\nu_0,\nu_1)+\varepsilon \tag{194}$$

We compute

$$H(\nu_i|m)=\sum_{j=1}^{n}\int_{A_j}a_{i.j}^{h}\log a_{i.j}^{h}dm=\sum_{j=1}^{n}a_{i.j}^{h}\log a_{i.j}^{h}m_h(\{x_j\})=H(\nu_i^h|m_h), \tag{195}$$

For $\mathrm{i}=0,1$. Denote $\eta_t^h(\{x_j\}):=\eta_t(A_j), j=1,2...,n$. Suppose $\eta_t=\rho_t.m$. From Jensen's inequality we get

$$H(\eta_t^h|m_h)=\sum_{i=1}^{\infty}\frac{\int_{A_j}\rho_t dm}{m(A_j)}\log\frac{\int_{A_j}\rho_t dm}{m(A_j)}m_h(\{x_j\})$$

$$\leq\sum_{j=1}^{\infty}\left(\frac{1}{m(A_j)}\int_{A_i}\rho t\log\rho t\ dm\right)m_h(\{x_i\})=H(\eta_t|m),$$

which together with (194) and (195) implies

$$H(\eta_t^h|m_h)\leq(1-t)H(\nu_0^h|m_h)+tH(\nu_1^h|m_h)-\frac{K}{2}t(1-t)W^2(\nu_0,\nu_1)+\varepsilon \tag{196}$$

Firstly, we consider the case $K<0$. Let q^h be $\mathrm{a}-2\,\mathrm{R}(\mathrm{h})$-optimal coupling of ν_0^h

and ν_1^h.Then the formula

$$\hat{q}:=\sum_{j,k=1}^{n}\left[q^h(\{(x_j,x_k)\})\delta_{x_j,x_k}\times\frac{1_{A_j}\times A_k}{m(A_j)m(A_k)}(m\times m)\right]$$

defines a measure on $M_h\times M_h\times M\times M$ which has marginals ν_0^h,ν_1^h and ν_1 Moreover, the projection of $\hat{q}$ on the first two factors is equal to q^h. Therefore we have

$$W(\nu_0,\nu_1)^2 \le \int d(x,y)^2 d\hat{q}(x^h,y^h,x,y)$$

$$\le \int \left[d(x.x^h)+d(x^h,y^h)+d(y^h,y)\right]^2 d\hat{q}(x^h,y^h,x,y)$$

$$=\sum_{j,k=1}^{n} \frac{\hat{q}(\{(x_j,x_k)\})}{m(A_j)m(A_k)} \int_{A_j\times A_k} \left[d(x,x_j)+d(x_j,x_k)+d(x_k,y)\right]^2 dm(x)(y)$$

$$\le \sum_{j,k=1}^{n} q^h(\{(x_j,x_k)\})\left(d(x_j,x_k)+2R(h)\right)^2 = W^{-2R(h)}(\nu_0^h,\nu_1^h)^2,$$

which together with (196) yields

$$H(\eta_t^h|m_h) \le (1-t)H(\nu_0^h|m_h)+tH(\nu_1^h|m_h)-\frac{K}{2}t(1-t)W^{-2R(h)}(\nu_0^h,\nu_1^h)+\varepsilon. \qquad (197)$$

In the case $K<0$ we start with an optimal coupling q of ν_0 and ν_1 and we show that the measure

$$q^{-h} := \sum_{j,k=1}^{n} q(A_j\times A_k)\delta_{(x_j,x_k)}$$

is a coupling of ν_0^h and ν_1^h. Indeed, if $A\subset M_h$ then we have in turn

$$\sum_{j,k=1}^{n} q(A_j\times A_k)\delta_{(x_j,x_k)}(A\times M_h) = \sum_{j,k=1}^{n} q(A_j\times A_k)\delta_{x_j}(A) = \sum_{j,k=1}^{n} q(A_j\times M)\delta_{x_j}(A)$$

$$= \sum_{j,k=1}^{n} v_0(A_j)\delta_{x_j}(A) = \sum_{j,k=1}^{n} v_0^h(x_j)\delta_{x_j}(A) = v_0^h(A)$$

Since for any $j,k=1,2....,n$ and for arbitrary $x\in A_j$ and $y\in A_k$ we have $\left(d(x_j,k_j)-2R(h)\right)_+ \le \left(d(x_j,x_k)-d(x,x_j)-d(y,x_k)\right)_+ \le d(x,y)$ one can estimate:

$$W^{+2R(h)}(\nu_0^h,\nu_1^h)^2 \le \sum_{j,k=1}^{n} q(A_j\times A_k)\left[\left(d(x_j,x_k)-2R(h)\right)_+\right]^2$$

$$= \sum_{j,k=1}^{n} \int_{A_j\times A_k} \left[\left(d(x_j,x_k)-2R(h)\right)+\right]^2 dq(x,y)$$

$$\le \sum_{j,k=1}^{n} \int_{A_j\times A_k} \left[\left(d(x_j,x_k)-d(x,x_j)-d(y,x_k)\right)+\right]^2 = dq(x,y)$$

$$\le \sum_{j,k=1}^{n} \int_{A_j\times A_k} d(x_j,x_k)^2 dq(x,y) = \int_{M\times M} d(x,y)^2 dq(x,y) = W(\nu_0,\nu_1)$$

Therefore from (196) we obtain

$$H\left(\eta_t^h|m_h\right) \le (1-t)H\left(V_0^h|m_h\right) + H\left(V_1^h|m_h\right) - \frac{K}{2}t(1-t)W^{+2R(h)}\left(V_0^h, V_1^h\right)^2 + \varepsilon \qquad (198)$$

For ε sufficiently small we can get

$$-\frac{K}{2}t(1-t)W^{+2R(h)}\left(V_0^h, V_1^h\right)^2 + \varepsilon \le -\frac{K}{2}t(1-t)W^{+2R(h)}\left(V_0^h, V_1^h\right)^2 \qquad (199)$$

and then (197), (198) yield

$$H\left(\eta_t^h|m_h\right) \le (1-t)H\left(V_0^h|m_h\right) + H\left(V_1^h|m_h\right) - \frac{K}{2}t(1-t)W^{\pm h}\left(V_0^h, V_1^h\right)^2 \qquad (200)$$

depending on the sign of K . The inequality (199) fails only when $K > 0$ and $W^{+h}\left(V_0^h, V_1^h\right) = 0,$ but in this case $W\left(V_0^h, V_1^h\right) \le h$ and either $\eta = V_0^h$ or $\eta = V_1^h$ verifies directly the condition (186) from the definition of h-rough curvature bound for the discretization.

The measure $\pi = \sum_{j=1}^{n} \eta_t^h\left(\{x_j\}\right)\delta_{x_j} \times 1_{A_j}\eta_t$ is a coupling of η_t^h and η_t, so

$$W^2\left(\eta_t^h, \eta_t\right) \le \int_{M_h \times M} d^2(x,y)\,d\pi(x,y) \le R^2(h)$$

and similarly $W^2\left(V_i^h, V_i\right) \le R^2(h)$. For $i = 1,2.$ Because η_t is an ε-rough t -approximate point between V_0 and V_1 we deduce

$$W\left(\eta_t^h, V_0^h\right) \le W\left(\eta_t, V_0\right) + 2R(h) \le tW\left(V_0, V_1\right) + 2R(h) + \varepsilon \le tW\left(V_0^h, V_1^h\right) + 2R(h)(1+t) + \varepsilon$$

and by a similar argument

$$W\left(\eta_t^h, V_1^h\right) \le (1-t)W\left(V_0^h, V_1^h\right) + 2R(h)(2-t) + \varepsilon$$

From (193) we conclude that η^h is an h -rough t -approximate point between V_0^h

and V_1^h ,which together with (200) proves that $h - Curv\left(M_h, d, m_h\right) \ge K.$

(iii) follows the same lines as (ii).

Example(4.3.11)[7]: If we consider on Z^n the metric d_1 coming from the norm $|\cdot|_1$ in R^n defined by $|r|_1 = \sum_{i=1}^{n} |x_i|$ and with the measure $\bar{m}_n = \sum_{x \in Z^n} \delta_x$ then $h - Curv\left(Z^n, d_1, \bar{m}_n\right) \ge 0$ for any $h \ge 2n$.

The n -dimensional grid E^n having Z^n as set of vertices, equipped with the graph distance and with the measure m_n which is the 1-dimensional Lebesgue measure on the edges, has $h-Curv\left(E^n,d_1,m_n\right)\geq 0$, for any $h\geq 2(n+1)$.

Proof. We use the following result:

Lemma (4.3.12)[7]: (See [22]). Any finite dimensional Banach space equipped with the Lebesgue measure has curvature ≥ 0.

We tile the space R^n with n -dimensional cubes of edge 1 centered in the vertices of the grid. The $|\cdot|_1$-radius of the cells of the tessellation with such cubes is $n/2$. Therefore, claim (i) is a consequence of Theorem (4.3.10) (iii) applied to the space $\left(R^n,|\cdot|_1\,dx\right)$, and of Lemma (4.3.12).

For the proof of (ii) we follow the same argument like in the proof of Theorem (4.3.10). In this case, we pass from a probability on the grid to a probability on R^n by averaging on each cube of the tessellation and scaling. Here one should take into account that for a cube C from the tiling

$$\sup\left\{|x-y|_1 : x\in C\bigcap E^n, y\in C\right\}=\frac{n+1}{2},$$

that provides the minimal $h=2(n+1)$ starting from which
$h-Curv\left(E^n,d_1,m_n\right)\geq 0$

Example (4.3.13)[7] (i): Let G be the graph that tiles the Euclidian plane with equilateral triangles of edge r . We endow G with the graph metric d_G induced by the Euclidian metric and with the 1-dimensional Lebesgue measure m on the edges. Then G has h -curvature ≥ 0 for any $h\geq 8r\sqrt{3}/3$

The graph G' that tiles the Euclidian plane with regular hexagons of edge length r , equipped as usual with the graph metric $d_{G'}$ and with the 1-dimensional measure m', has h –curvature ≥ 0 for any $h\geq 34r/3$..

Proof: Consider a Cartesian coordinate system in the Euclidian plane with origin O and axes Ox and Oy . We equip R^2 with the Banach norm $\|\cdot\|$ ·that has as unit ball the regular hexagon centered in O , having two opposite vertices on Ox and the edge length (measured in the Euclidian metric) equal to 1. Explicitly

$$\|(x,y)\|=\max\left\{\frac{3\sqrt{3}}{3}|y|,|x|+\frac{\sqrt{3}}{3}|y|\right\}$$

for any (x,y) in R^2 . We denote by d the metric determined by this norm.

For the triangular tessellation we choose the origin O to be one of the vertices of the graph and two of the 6 edges emanating from O be along the Ox axis. The edges of the graph have length r in the Euclidian metric. We see that

$d_G(V_1,V_2)=d(V_1,V_2)$ for any two vertices v_1 and v_2 of the graph. In general for $x,y\in G$ we have $d_G(x,y)-d(x,y)\le r$. Then one can construct a coupling ˆ $\hat{d}$ of d_G and d by setting $\hat{d}(V,x):=d(V,x)$ for v vertex of G and $x\in R^2$ and $\hat{d}(y,x):=\inf_{i=1,2}\{d_G(y,V_i)+d(V_i,x)\}$

if $y\in G$ belongs to an edge with endpoints V_1,V_2 and $x\in R^2$

By Lemma (4.3.12) $Curv(R^2,d,\lambda)\ge 0$ where λ is the 2-dimensional Lebesgue measure. If we tile the plane with regular hexagons $A_j, j\in N$, which have vertices in the centers of the triangles $\hat{d}(y,x)\le 2r\sqrt{3}/3$ for any $y\in A_j\cap G$ and $x\in A_j$.The proof of the h -curvature bound is a modification of the proof of Theorem (4.3.10). We start with $V_0,V_1\in P_2^*(G,d_G,m)$ with $V_i=\rho_i m, i=0,1$, and we define

$$\bar{V}:=\sum_{j=1}^{\infty}\frac{1}{\lambda(A_j)}\left(\int_{G\cap A_j}\rho_i dm\right)1_{A_j}\lambda\in P_2^*\left(R^2,d,\lambda\right)\text{for } i=0,1$$

We have then $\hat{d}w(V_i,\bar{V}_i)\le 2r\sqrt{3}/3$. We consider $\tilde{\eta}_t=\tilde{\rho}_t.\lambda$ the geodesic that joints $\tilde{V}_0$ and $\tilde{V}_1$, along which the convexity condition for the entropy on $P_2^*(R^2,d,\lambda)$ is fulfilled and denote

$$\eta_t:=\sum_{j=1}^{\infty}\frac{1}{m(G\cap A_j)}\left(\int_{A_j}\tilde{\rho}_t d\lambda\right)1_{G\cap A_j}.m$$

Then η_t is $8r\sqrt{3}/3-$rough t -approximate point between V_0 and V_1. From Jensen's inequality we obtain $H(\eta_t|m)H(\tilde{\eta}_t|\lambda)-\log m(G\cap A)+\log\lambda(A)$ and $H(\tilde{V}_i|\lambda)\le H(V_i|m)+\log m(G\cap A)-\log\lambda(A)$ observe that all sets A_j have the same Lebesgue measure $\lambda(A)$ and all sets $G\cap A_j$ have the same measure $m(G\cap A)$. Hence η_t satisfies

$$H(\eta_t|m)\le(1-t)H(V_0|m)+tH(V_1|m)$$

and so we have proved $h-Curv(G,d_G,m)\ge 0$ for any $h\ge 8r\sqrt{3}/3$.

(ii) For the hexagonal tessellation let O be again one of the vertices of the graph and one of the 3 edges emanating from it be along the Oy axis. In this case we use the Banach norm $\|\cdot\|' := \frac{3}{4}\|\cdot\| \cdot$ on R^2 and denote by d' the associated metric. The length of the edges of the graph in the metric d' is equal to $4r/3$. We see that $|d_G - d'| \le r/3$ for any two vertices V_1, V_2 with $d_G(V_1, V_2) = 2kr,\ K \in N$. In general $|d_G - d'| \le r/3$ on the set of vertices and $|d_G - d'| \le r$ everywhere on G'

One can construct then a coupling $\hat{d}'$ of $d_{G'}$ and d' in the following way: Fix

$V_0 = O$ .If V is a vertex of the graph with $d_{G'}(V_0, V_1) = 2kr,\ k \in N$ then set $\hat{d}'(V, x) := d'(V, x),\ x \in R^2$.For $y \in G'$ with $d_G(V_0, V) \ne 2kr$, $k \in N$ define

$$\hat{d}'(y, x) := \inf\{d_{G'}(y, V) + d'(V, x) : V \in G', d_{G'}(V_0, V) = 2kr\}$$

We tile the plane with equilateral triangles $B_i, i \in N$, with vertices in the centers of the hexagons of the graph. Then $\hat{}\hat{d}'(y, x) \le 17r/6$ for $y \in B_i \cap G',\ x \in B_i$. By the same argument as for the triangular tiling we obtain $h - Curv(G', d_{G'}, m') \ge 0$ for any

$$h \ge 4.17r/6 = 34r/3.$$

The following result is probably well-known.

Lemma (4.3.14)[7]:(i) *If* $\frac{1}{l} + \frac{1}{n} < \frac{1}{2}$ then $G(l, n, r)$ can be embedded into the 2-dimensional hyperbolic space with constant sectional curvature

$$K = -\frac{1}{r^2}\left[\operatorname{arccos} h\left(2\frac{\cos^2\frac{\pi}{n}}{\sin^2\frac{\pi}{l}} - 1\right)\right]^2 \qquad (201)$$

There are infinitely many choices of such l and n. In any case, the graph is unbounded.

(ii) If $\frac{1}{l} + \frac{1}{n} > \frac{1}{2}$ then $G(l, n, r)$ is one of the five regular polyhedra (Tetrahedron, Octahedron, Cube , Icosahedrons , Dodecahedron) and can be embedded into the

2 -dimensional sphere with constant sectional curvature

$$K=-\frac{1}{r^2}\left[\arccos\left(2\frac{\cos^2\left(\frac{\pi}{n}\right)}{\sin^2\left(\frac{\pi}{l}\right)}-1\right)\right]^2 \qquad (202)$$

If $\frac{1}{l}+\frac{1}{n}=\frac{1}{2}$ then $G(l,n,r)$ can be embedded into the Euclidian plane $(\mathrm{K}=0)$. In this case there are exactly three cases corresponding to the 3 regular tessellations of the Euclidian plane: the tessellation of triangles $(l=6, n=3)$, of squares $(l=n=4)$, and of hexagons $(l=3, n-6)$

Proof: Firstly we see that

$$2\frac{\cos^2\frac{\pi}{n}}{\sin^2\frac{\pi}{l}}-1>1\Leftrightarrow \sin^2\left(\frac{\pi}{2}-\frac{\pi}{n}\right)>\sin^2\left(\frac{\pi}{l}\right)\Leftrightarrow\frac{1}{l}+\frac{1}{n}<\frac{1}{2}$$

hence in each case the expression that defines the curvature K makes sense.

For given l,n,r we construct the embedding in the following way: we start from an arbitrary point O of the 2-hyperbolic space with curvature K, denoted by $H^{k,2}$. From this point we construct n geodesic lines $OA_1, OA_2, ..., OA_n$ of length

$$R:=\frac{1}{\sqrt{-K}}\operatorname{arcsin}h\left(\frac{\sinh\sqrt{-K}r}{\sin\left(\frac{2\pi}{n}\right)}\sin\left(\frac{\pi}{l}\right)\right) \qquad (203)$$

such that the inner angle between any two consecutive geodesics OA_k, OA_{k+2} is $2\pi/n$. We prove that $A_1, A_2...., A_n$ correspond to vertices of the given graph, and the geodesics $A_1A_2,...,A_{n-1}A_nA_1$ correspond isometrically to consecutive edges in $G(l,n,r)$ that bound a regular n-polygon with edge-length r and all angles equal to $2\pi/l$. Let us denote by

d the intrinsic metric on $H^{k,2}$.

From the Cosine Rule for hyperbolic triangles applied to ΔOA_1A_2 and from (201) and (203) we have:

$$\cosh\left(\sqrt{-k}\,d\left(A_1,A_2\right)\right)=\cosh^2\left(\sqrt{-k}R\right)-\sin\mathrm{h}^2\left(\sqrt{-k}R\right)\cos\left(\frac{2\pi}{n}\right)$$

$$= 1 + \sin h^2\left(\sqrt{-kR}\right)\left(1-\cos\left(\frac{2\pi}{n}\right)\right) \quad = 1 + \frac{\sin h^2\left(\sqrt{-kR}\right)}{\sin^2\left(\frac{2\pi}{n}\right)}\sin^2\frac{\pi}{l}\left(1-\cos\frac{2\pi}{n}\right)$$

$$= 1 + \frac{\sin h^2\left(\sqrt{-kR}\right)-1}{1+\cos\frac{2\pi}{n}}\sin^2\frac{\pi}{l} = 1 + \frac{\sin h^2\frac{\pi}{l}}{2\cos^2\frac{\pi}{n}}\left[\left(2\frac{\cos^2\frac{\pi}{n}}{\sin^2\frac{\pi}{l}}-1\right)^2-1\right]$$

$$= \frac{\cos^2\frac{\pi}{n}}{\sin^2\frac{\pi}{l}} - 1 = \cosh\left(\sqrt{-k}\,r\right)$$

So $d(A_1, A_2) = r$ and the same holds for all the other edges of the polygon. We apply now the Sine Rule for the hyperbolic triangle ΔOA_1A_2 and (203) in order to compute:

$$\sin \mathrm{S}\left(A_1; O, A_2\right) = \frac{\sin\left(\frac{2\pi}{n}\right)}{\sinh\sqrt{-kr}}\sinh\sqrt{-k}R = \sin\left(\frac{\pi}{l}\right) \tag{204}$$

where $\mathrm{S}\left(A_1; O, A_2\right)$ denotes the angle at A_1 in the triangle ΔOA_1A_2. This angle is less than $\pi/2$ because it is equal to $\mathrm{S}\left(A_2; O, A_1\right)$ and in the hyperbolic triangles the sum of the angles of a triangle is less than π . Therefore (204) shows that all the angles of the polygon are equal to $2\pi/l$, so around each vertex one can construct other $l-1$

polygons with n edges, congruent with the first one.We repeat the procedure with each of the vertices of the new polygons. In this way the whole space $H^{k,2}$ can be tiled with regular polygons which are faces of the graph $G(l,n,r)$

(ii), (iii) Since there is only a finite number of examples with well-known realizations, the claim can be verified directly. Alternatively, one can prove it like in the part (i) with appropriate interpretations of the hyperbolic sine as sine for positive curvature and as length for the Euclidian plane.

Theorem(4.3.15)[7]: For any numbers $l, n \geq 3$ and for any $r > 0$ both metric measure spaces $V(l,n,r), d.\tilde{m}$ and $\left(G(l,n,r), d.m\right)$ have h-curvature $\geq K$ for $h \geq r.C(l,n)$ where

$$K=\begin{cases} -\frac{1}{r^2}\left[\operatorname{arccos}h\left(2\frac{\cos^2\left(\frac{\pi}{n}\right)}{\sin^2\left(\frac{\pi}{l}\right)}-1\right)\right]^2 & for\ \frac{1}{l}+\frac{1}{n}>\frac{1}{2} \\ -\frac{1}{r^2}\left[\arccos\left(2\frac{\cos^2\left(\frac{\pi}{n}\right)}{\sin^2\left(\frac{\pi}{l}\right)}-1\right)\right]^2 & for\ \frac{1}{l}+\frac{1}{n}<\frac{1}{2} \\ 0 & for\ \frac{1}{l}+\frac{1}{n}<\frac{1}{2} \end{cases} \tag{205}$$

and

$$C(l,n)=4.\operatorname{arcsin}h\left(\frac{1}{\sin\left(\frac{\pi}{n}\right)}\sqrt{\frac{\cos^2\left(\frac{\pi}{n}\right)}{\sin^2\left(\frac{\pi}{l}\right)}-1}\right)\Big/\operatorname{arccos}h\left(2\frac{\cos^2\left(\frac{\pi}{n}\right)}{\sin^2\left(\frac{\pi}{l}\right)}-1\right)$$

Proof : We look at $\mathrm{V}(l,n,r)$ and $G(l,n,r)$ as subsets of the 2-manifold with constant curvature K(given by Lemma (4.3.14). We tile the manifold with the faces of the dual graph $G(l,n,r')$having vertices in the centers of the faces of $G(l,n,r)$ the center O of the polygon with n edges in the proof of Lemma (4.3.14). becomes vertex of the dual).

We make explicitly the calculations only in the hyperbolic case, the other two cases are similar. One can decompose the hyperbolic space as $H^{k,2}=\bigcup_{j=1}^{\infty}F_j$ where $\{F_j\}_j$ are the faces of the dual graph, as described above. The curvature bound for the discrete space $V(l,n,r)$is then a consequence of the Theorem (4.3.10). For $G\doteq G(l,n,r)$

the proof of the curvature bound is a modification of the proof of Theorem(4.3.9) We start with $\mathrm{V}_0,\mathrm{V}_1\in\mathrm{P}_2^*\left(G(l,n,r),d,m\right)$ with $\mathrm{V}_i=\rho_i.m,i=0,1$

$\bar{\mathrm{V}}_i:=\sum_{j=1}^{\infty}\frac{1}{vol\left(F_j\right)}\left(\int_{G\cap F_j}\rho_i dm\right)1_{F_j}.vol\in\mathrm{P}_2^*\left(H^{k,2},d,vol\right)$ for $i=0,1$

Now the place of$R(h)$ from Theorem(4.3.10) is taken by R from the proof of Lemma (4.3.14)(i), so $\mathrm{W}\left(\mathrm{V}_i,\tilde{\mathrm{V}}_i\right)\le R$ One can express R only in terms of our initial data l,n and r as $R = rC(l,n)/4$, with $C(l,n)$ given in the statement of the theorem. We consider $\tilde{\eta}_t = \tilde{\rho}_t.\text{vol}$ the geodesic that joints $\tilde{V}_0$ and $\tilde{V}_1$ along which one has the K-convexity for the entropy on $H^{k,2}$ [141] and denote .

$$\eta_t := \sum_{j=1}^{\infty} \frac{1}{m\left(G \cap F_j\right)} \left(\int_{F_j} \tilde{\rho}_t dvol \right) 1_{G \cap F_j}.m.$$

Then η_t is $4R$ – rough t-approximate point between V_0 and V_i. From Jensen's inequality we obtain $H(\eta_t|m) \leq H(\bar{\eta}_t|vol) - \log m(G \cap F) + \log vol(F)$ and $H(\tilde{V}_t|vol) \leq H(V_i|m) + \log m(G \cap F) - \log vol(F)$ observe that all faces F_j have the same volume $vol(F)$ and all sets $G \cap F_j$ have the same measure $m(G \cap F)$.
Hence, like in the proof of Theorem (4.3.10) η_t satisfies so we have proved $h - Curv(G(l,n,r),d,m) \geq k$ for any $h \geq 4R$ in the hyperbolic axe $(k < 0)$

In [270] the combinatorial curvature of a graph G is a map $\Phi_G : V(G) \to \mathbf{R}$ that assigns to each vertex $x \in V(G)$ the number $\Phi_G(x) = 1 - \frac{m(x)}{2} + \sum_{i=1}^{m(x)} \frac{1}{d(F_i)}$ where $m(x)$ is the degree of the vertex $x, d(F)$ is the number of edges of the cycle bounding a face F , and $F_1, F_2 ....., F_{m(x)}$ are the faces around the vertex x The combinatorial curvature introduced in [164] is a map $\Phi_G^* : F(G) \to \mathbf{R}$, where the curvature $\Phi_G^*(F)$ of a face F is given by the curvature Φ_G of the corresponding vertex in the dual graph. For the homogeneous graph $G(l,n,r)$ the curvature of any vertex x is $\Phi_G(x) = l\left(\frac{1}{l} + \frac{1}{n} - \frac{1}{2}\right)$ and the curvature in the sense of Gromov [164] of any face F is $\Phi_G^*(F) = n\left(\frac{1}{l} + \frac{1}{n} - \frac{1}{2}\right)$.

Note that the sign of the combinatorial curvature in both approaches above changes according to whether $\frac{1}{l} + \frac{1}{n}$ is greater or less than $\frac{1}{2}$ Rather curiously, in our Theorem (4.3.15) the sign of the rough curvature bound changes in the same manner, although our notion of curvature applies to graphs that have a metric structure and a reference measure. For the moment we see no further links with the notions of combinatorial curvature mentioned here.

Let (M,d) be a metric space and $m \in \mathrm{P}_2(M,d)$ be a given probability measure. The measure m is said to satisfy a Talagrand inequality (or a transportation cost inequality) with constant *Kiff* for all $\mathrm{V} \in \mathrm{P}_2(M,d)$

$$\mathrm{W}(\mathrm{V},m) \le \sqrt{\frac{2H(\mathrm{V}|m)}{K}} \tag{206}$$

Such an inequality was first proved by Talagrand in [179] for the canonical Gaussian measure on R^2 . A positive rough curvature bound allows us to obtain a weaker inequality, in terms of the perturbation W^{+h} of the Wasserstein distance:

Proposition (4.3.16)[7]: ("h-Talagrand inequality"). Assume that (M,d,m) is a metric measure space which has $h-Curv(m,d,m) \ge K$ for some numbers $\mathrm{h} > 0$ and $\mathrm{K} > 0$. Then for each $\mathrm{V} \in \mathrm{P}_2(M,d)$ we have

$$\mathrm{W}^{+h}(\mathrm{V},m) \le \sqrt{\frac{2H(\mathrm{V}|m)}{K}} \tag{207}$$

We will call (207) $h-Talagrand$ inequality.

Proof: Since we assumed that m is a probability measure, for any $\mathrm{V} \in \mathrm{P}_2(M,d)$

the entropy functional is nonnegative: $H(\mathrm{V}|m) \ge -\log m(M) = 0,$, according to [7]:The curvature bound $h-Curv(M,d,m) \ge K$ implies that for the pair of measures ν and m and for each $t \in [0,1]$ there exists an h -rough t - approximate point $\eta_t \in \mathrm{P}_2(M,d)$ such that

$$H(\eta_t|m) \le (1-t)H(\mathrm{V}|m) - \frac{K}{2}t(1-t)\mathrm{W}^{+h}(\mathrm{V}|m)^2 \tag{208}$$

If $H(\mathrm{V}|m) < \frac{K}{2}\mathrm{W}^{+h}(\mathrm{V},m)$ then there exists an $\varepsilon > 0$ such that $H(\mathrm{V}|m) + \varepsilon < \frac{K}{2}\mathrm{W}^{+h}(\mathrm{V},m)^2$ This together with (208) would imply

$$H(\eta_t|m) < \frac{K}{2}t(1-t)\mathrm{W}^{+h}(\mathrm{V}|m)^2 - \varepsilon(1-t)$$

for each $t \in [0,1]$. We choose now t very close to 1, such that $0 < 1-t < \varepsilon$ and

$K(1-t)^2 W^{-h}(V,m)^2 < \varepsilon^2$. This entails $H(\eta_t|m) < -\varepsilon^2/2 < 0$ in contradiction with the fact that the entropy functional is nonnegative. Therefore $H(V|m) \geq -\varepsilon^2 \frac{K}{2} W^{+h}(V,m)^2$, which is precisely our claim.

A Talagrand inequality for the measure m implies a concentration of measure inequality for m (see for instance [136]).

For a given Borel set $A \subset M$ denote the (open) $r-$*neighborhood* of A by $B_r(A) := \{x \in M : d(x,A) < r\}$ for $r > 0$. The concentration function of (M,d,m) is defined as $a_{(M,d,m)}(r) := \sup\left\{1 - m(B_r(A)) : A \in B(M), m(A) \geq \frac{1}{2}\right\}, r > 0$

We refer to [172] for further details on measure concentration.

The following result shows that positive rough curvature bound implies a normal concentration inequality, via h-Talagrand inequality.

Proposition(4.3.17)[7]: Let (M,d,m) be a metric measure space with $h-\text{C}urv(M,d,m) \geq K > 0$ for some h>0. Then there exists an $r_0 > 0$ such that for all $r > r_0$

$$a_{(M,d,m)}(r) \leq e^{-Kr^2/8}$$

Proof: We follow essentially the argument of K. Marton used in [9] for obtaining concentration of measure out of a Talagrand inequality for the Wasserstein distance of order 1. Let $A, B \in B(M)$ be given with $m(A).m(B) > 0$ Consider the conditional probabilities $m_A = m(\cdot|A)$ and $m_B = m(\cdot|B)$. For these measures the h -Talagrand inequality holds:

$$W^{+h}(m_A.m) \leq \sqrt{\frac{2H(m_A|m)}{K}}, W^{+h}(m_B.m) \leq \sqrt{\frac{2H(m_B|m)}{K}} \tag{209}$$

Let q_A and q_B be the $+\text{h}-$optimal couplings of m_A, m and m_B, m respectively. According to [2], there exists a probability measure $\hat{q}$ on $M \times M \times M$ such that its projection on the first two factors is q_A and the projection on the last two factors is q_B. Then we have in turn

$$W^{+h}(m_A,m) + W^{+h}(m_B,m) = \left\{\int_{M\times M\times M} \left[(d(x_1,x_2)-h)_+\right]^2 d\hat{q}(x_1,x_2,x_2)\right\}^{\frac{1}{2}}$$

$$+\left\{\int_{M\times M\times M}\left[\left(d\left(x_2,x_3\right)-h\right)_+\right]^2 d\hat{q}\left(x_1,x_2,x_2\right)\right\}^{\frac{1}{2}}$$

$$\geq\left\{\int_{M\times M\times M}\left[\left(d\left(x_1,x_2\right)-h\right)_+ + d\left(x_2,x_3-h\right)_+\right]^2 d\hat{q}\left(x_1,x_2,x_2\right)\right\}^{\frac{1}{2}}$$

$$\geq\left\{\int_{M\times M\times M}\left[\left(d\left(x_1,x_2\right)+d\left(x_2,x_3\right)-2h\right)_+\right]^2 d\hat{q}\left(x_1,x_2,x_2\right)\right\}^{\frac{1}{2}}$$

$$\geq\left\{\int_{M\times M\times M}\left[\left(d\left(x_1,x_3\right)-2h\right)_+\right]^2 d\hat{q}\left(x_1,x_2,x_2\right)\right\}^{\frac{1}{2}}$$

Assume now that $d(A,B)\geq 2h$. Since the projection on the first factor of $\hat{q}$ is m_A and the projection on the last factor is m_B , the support of $\hat{q}$ must be a subset of $A\times M\times B$,

hence

$$\left\{\int_{M\times M\times M}\left[\left(d\left(x_1,x_3\right)-2h\right)_+\right]^2 d\hat{q}\left(x_1,x_2,x_2\right)\right\}^{\frac{1}{2}}\geq d\left(A,B\right)-2h$$

The above estimates together with (209) imply

$$d\left(A,B\right)-2h\leq\sqrt{\frac{2H\left(m_A|m\right)}{K}}+\sqrt{\frac{2H\left(m_B|m\right)}{K}}=\sqrt{\frac{2}{K}\log\frac{1}{m\left(A\right)}}+\sqrt{\frac{2}{K}\log\frac{1}{m\left(B\right)}}$$

if we choose now $2h\leq r$ and for a given $A\in B(M)$ we replace B by $\complement B_r(A)$, we get

$$r-2h\leq\sqrt{\frac{2}{K}\log\frac{1}{m(A)}}+\sqrt{\frac{2}{K}\log\frac{1}{1-m\left(B_r\left(A\right)\right)}}$$

Hence, for $m(A)\geq\frac{1}{2}$,

$$r-2h\leq\sqrt{\frac{2}{K}\log 2}+\sqrt{\frac{2}{K}\log\frac{1}{1-m\left(B_r\left(A\right)\right)}}$$

Therefore whenever $r\geq 2\sqrt{\frac{2}{K}\log 2}+4h$ for instance we have

$$\frac{r}{2}\leq\sqrt{\frac{2}{K}\log\frac{1}{1-m\left(B_r\left(A\right)\right)}}$$

or equivalently

$$1-m\left(B_r\left(A\right)\right)\le e^{-Kr^2/8}$$

which ends the proof.

In [254] it has been shown that a Talagrand type inequality implies exponential integrability of the Lipshitz functions. We prove further that an h – Talagrand inequality leads to the same conclusion.

Theorem (4.3.18)[7]: Assume that (M,d) is a metric space and let $h>0$ be given. If m is a probability measure on (M,d)that satisfies an h – Talagrand inequality of constant $K>0$ then all Lipschitz functions are exponentially integrable. More precisely, for any Lipschitz function ϕ. with. $\|\phi\|_{Lip}\le 1$and $\int \phi dm=0$ we have

$$\int_M e^{t\phi}dm\le e^{\frac{t^2}{2k}+h}\qquad \forall t>0 \tag{210}$$

or equivalently, for any Lipschitz function ϕ.

$$\int_M e^{t\phi}dm\le \exp\left(t\int_M \phi dm\right)\exp\left(\frac{t^2}{2k}\|\phi\|_{Lip}^2\ ht\|\phi\|_{Lip}\right).$$

(211)

Proof: The proof we present here extends the one given in [54]. Let f be a probability density with f log f integrable w. r. t m .The -Talagrand inequality implies

$$\mathrm{W}^{+h}\left(fm,m\right)\le\sqrt{\frac{2}{k}\int_M f\log fdm}\le\frac{t}{2k}+\frac{1}{t}\int_M f\log fdm$$

for each $t>0$.We consider now the Wasserstein distance of order 1 of two probability measures μ and V

$$\mathrm{W}^1\left(\mu,\mathrm{V}\right):=\inf\int_{M\times M}d\left(x_{0,}x_1\right)dq\left(x_{0,}x_1\right),$$

where q ranges over all couplings of μ and V .If $\tilde{q}$ is a +h-optimal coupling of fm and m then by the Cauchy–Schwartz inequality,

$$\mathrm{W}^{+h}\left(fm,m\right)=\left\{\int_{M\times M}\left[\left(d\left(x_{0,}x_1\right)-h\right)_+\right]^2 d\tilde{q}\left(x_{0,}x_1\right),\right\}^{1/2}$$

$$\geq \int_{M\times M}\left[\left(d\left(x_{0,}x_1\right)-h\right)_+\right]^2 d\tilde{q}\left(x_{0,}x_1\right)\geq \mathrm{W}^1\left(fm,m\right)-h$$

The Kantorovich–Rubinstein theorem gives the following duality formula

$$\mathrm{W}^1\left(fm,m\right)=\sup_{\|\phi\|_{Lip\leq 1}}\left\{\int_M \phi\, dm . \int_M \phi\, dm\right\}.$$

If ϕ is a Lipschitz function that satisfies the assumptions of the theorem ($\|\phi\|_{Lip}\leq 1$ and $\int \phi dm = 0$)then

$$\int_M \phi f dm \leq \mathrm{W}^{+h}\left(fm,m\right)+h\leq \frac{t}{2k}+\frac{1}{t}\int f\log f dm + h$$

which can be written as

$$\int_M\left(t\phi-\frac{t}{2k}\right)fdm\leq \int_M f\log f dm + ht \tag{212}$$

This estimate should take place for any probability density f Therefore one can take

$$f=e^{t\varphi-\frac{t^2}{2k}}\left(\int_M e^{t\varphi-\frac{t^2}{2k}}dm\right)^{-1}$$

in formula (212) and obtain

$$\left\{\int_M\left(t\phi-\frac{t^2}{2k}\right)e^{t\phi-\frac{t^2}{2k}}dm\right\}\left(\int_M e^{t\phi-\frac{t^2}{2k}}\right)^{-1}$$

$$\leq \int_M e^{t\phi-\frac{t^2}{2k}}\left(\int_M e^{t\phi-\frac{t^2}{2k}}dm\right)^{-1}\left\{t\phi-\frac{t^2}{2k}-\log\left(\int_M e^{t\phi-\frac{t^2}{2k}}dm\right)\right\}dm+ht$$

This yields

$$\log\left(\int_M e^{t\phi-\frac{t^2}{2k}}dm\right)\leq ht,$$

that proves the claim (210). The general estimate (211) is a consequence of (210) applied to the function $\Psi=\frac{1}{\|\phi\|_{Lip}}\left[\phi-\int\phi dm\right]$.

Chapter 5

Dimensions and Relations Between Combinatorial Measurement

These relationships further extend and sharpen prior results concerning extensions of the Littlewood $2n/(n+1)$ -inequalities, the n-dimensional Khintchin inequalities, and the Kahane-Khintchin inequality. We show an estimate between the combinatorial structure of a series of $\alpha - Orlicz$ functions, that is finite and summation of norms of random variables in a Hilbert space.

Section (5.1): Dimension and Norms in Harmonic Analysis

The purpose of this section is to study a parameter that we call 'combinatorial dimension'; its definition has been motivated by previous work [228] where appropriate constructions in a harmonic analytic framework filled 'combinatorial' and 'analytic' gaps left open between Cartesian products of spectral sets.

We start with a set E (apriori devoid of structure), and a positive integer L. As us E^L denotes the L-fold Cartesian product of E,

$$\mathrm{E}^L = \{(x_1,...,x_L) : x_1,...,x_L \in \mathrm{E}\}.$$

Let F be an arbitrary subset of E^L and define for every positive integer s

$$\Psi_F(s) = \max\{|F \cap (A_1 \times ... \times A_L)| : A_1,...,A_L \subset \mathrm{E}, |A_1| = ... = |A_L| = s\} \qquad (1)$$

($|\cdot|$ denotes cardinality).

Definition (5.1.1)[226]: The combinatorial dimension of $F \subset \mathrm{E}^L$ is

$$\dim F = \inf\left\{a : \overline{\lim_{s\to\infty}} \frac{\Psi_F(s)}{s^a} < \infty\right\}.$$

$\dim F$ is exact if

$$\overline{\lim_{s\to\infty}} \frac{\Psi_F(s)}{s^{\dim F}} < \infty;$$

otherwise, dim F is asymptotic.

Next, we consider $\{\beta_n\}_{n\in N} S$, the Steinhaus system of (statistically) independent random variables equidistributed on the unit circle [127]. S is concretely realized as a sequence of functions defined on the probability space $\Omega = \otimes T$ where $T = [0, 2\pi)$ (with the usual Borel structure and Lebesgue measure) and $\otimes T$ is the direct product of T: For $w = \left(w(n)\right)_{n \in N} \in \Omega, \beta_0(w) \equiv 1$ and $\beta_n(w) = e^{iw(n)}, \quad n \geq 1.$

Taking into account the usual group structure on $[0, 2\pi)$, we shall view S as a set of algebraically independent characters on the compact abelian group Ω whose discrete dual group is $\oplus Z$,, where Z is the additive group of integers.

In this sections focusing on what we consider basic issues, we shall work in the framework of L-fold Cartesian products of S consisting of functions on Ω^L,

$$S^L = \left\{ (\beta_1, ..., \beta_L) : \beta_1, ..., \beta_L \in S \right\}$$

where

$$(\beta_1, ..., \beta_L)(w_1, ..., w_L) = \beta_1(w_1) ..., \beta_L(w_L).$$

Here we link the measurement of certain probabilistic-harmonic analytic properties of $F \subset S^L$ to the measurement of the combinatorial dimension of F; the analytic-combinatorial connections are summarized see [228] [231],{24],[267],[228] and[229].

We recall the definition of the $\Lambda(p)$ constant of $F \subset \Gamma, 2 < p < \infty$:

$$\lambda_F(p) = \sup \left\{ \frac{\|g\|_p}{\|g\|_2} : g \in L_F^2(G), g \neq 0 \right\},$$

where $L_F^2(G)$ denotes the space of L^2 functions on G whose spectrum is a subset of F.

Theorem(5.1.2)[226]: Let $F \subset S^L$ be arbitrary. For every integer $s > 0$

$$8^{-L}\left(\Psi_F(s)\right)^{1/2} \leq \lambda_F(2s) \leq \left(\Psi_F(s)\right)^{1/2}. \qquad (2)$$

Proof : We will denote the L canonical projections from S^L into S by $\pi_1, ..., \pi_L$:

$$\pi_i(\beta_1, ..., \beta_L) = \beta_i, \qquad 1 \leq i \leq L$$

We establish first the right hand inequality in (2). Let $s > 1$ be an arbitrary integer, and let

$$g = \sum_{\beta \in F} a_\beta \left(\pi_1(\beta), ..., \pi(\beta) \right)$$

be an arbitrary function in $L_F^2\left(\Omega^L\right)$. Write

$$g^s = \sum_{\beta_1 ... \beta_s \in F} a_{\beta_1} ... a_{\beta_s} \left(\pi_1(\beta_1) .... \pi_1(\beta_s), ..., \pi_L(\beta_1) ... \pi_L(\beta_s) \right).$$

with the aim of estimating

$$\|g\|_{2s}^{2s} = \|g^s\|_2^2.$$

To this end, observe first that for any $\gamma \in \Omega^L$

$$\left(g^s\right)^\wedge(\gamma) = \sum_{(\beta_1 ... \beta_s) \in A(\gamma)} a_{\beta_1} .. a_{\beta_s},$$

where

$$A(\gamma) = \left\{ (\beta_1, ..., \beta_s) \in F^s : \gamma = \left(\pi_1(\beta_1) ... \pi_1(\beta_s), ..., \pi_L(\beta_1) ... \pi_L(\beta_s) \right) \right\}$$

Therefore, by Schwartz's inequality,

$$\|g^s\|_2^2 = \sum_{\gamma \in \Omega^L} \left| \sum_{A(\gamma)} a_{\beta_1} ... a_{\beta_s} \right|^2 \leq \sum_{\gamma \in \Omega^L} |A(\gamma)| \sum_{A(\gamma)} A(\gamma) \left| a_{\beta_1} ... a_{\beta_s} \right|^2 \qquad (3)$$

Next, for any $\gamma \in \hat{\Omega}^L$, we estimate $|A(\gamma)| = 0$ as follows: Note that either $|A(\gamma)| = 0$ there exist

$\beta_1, ..., \beta_s \in F$

so that

$$\gamma = (\gamma_1, ..., \gamma_L) = \left(\pi_1(\beta_1) ... \pi_1(\beta_s), ..., \pi_L(\beta_1) ... \pi_L(\beta_s) \right), \qquad (4)$$

We assume (4). Now, observe that the algebraic independence of $S \subset \Omega$ implies that we have

$$(\gamma_1, ..., \gamma_L) = \gamma = \beta_1' ... \beta_s' \qquad (5)$$

for some $\beta_1' ... \beta_s' \in F$ only if

$$\begin{aligned}\{\pi_1(\beta_1'),...,\pi_1(\beta_s')\} &= \{\pi_1(\beta_1'),....,\pi_1(\beta_s')\} \equiv A_1\\ \vdots \qquad \vdots \qquad & \quad \vdots \qquad \vdots \qquad \vdots\\ \{\pi_L(\beta_1'),...,\pi_L(\beta_s')\} &= \{\pi_L(\beta_1),....,\pi_L(\beta_s)\} \equiv A_L\end{aligned}$$

That is, a necessary condition for: $\beta \in F$ to appear as a factor on the right hand side of (5) is that

$\pi_1(\beta) \in A_1,...\pi_L(\beta) \in A_L$

Therefore,

$$A(\gamma) \subset \{\beta \in F : \pi_1(\beta) \in A_1,....,\pi_L(\beta) \in A_L\}^s \tag{6}$$

Finally, since $|A_i| \le s$ for $i = 1,....,L,$ it follows from (6) and the definition of Ψ_F that

$$|A(\gamma)| \le (\Psi_F(s))^s. \tag{7}$$

Substituting (8) in (4), we obtain

$$\|g\|_{2s} = \|g'\|_2^{1/s} \le (\Psi_F(s))^{1/2} \left[\sum_{\gamma \in \Omega^L} \left(\sum_{A(\gamma)} |a_{\beta_1}...a_{\beta_s}|^2 \right) \right]^{1/2s}.$$

Clearly, $A(\gamma) \cap A(\gamma') = \phi$ whenever $\gamma \ne \gamma'$. Therefore,

$$\sum_{\gamma \in \Omega^L} \sum_{A(\gamma)} |a_{\beta_1}...a_{\beta_s}|^2 = \sum_{\beta_1,....,\beta_s \in F} |a_{\beta_1}|^2 ... |a_{\beta_s}|^2 = \left(\sum_{A(\gamma)} |a_\beta|^2 \right)^s$$

and the right hand side of (6) is established.

We now prove the left hand inequality in (2). Let s > 1 be arbitrary, and choose $A_1,....,A_L \subset S,$ where

$|A_1| = ... = |A|_L = s$

and

$$\Psi_F(s) = |F \cap (A_1 \times .... \times A_L)|. \tag{8}$$

Write the Riesz product

$$\mathbf{R}(w_1,...,w_L) = \left[\prod_{\beta \in A_1} (1 + \cos\beta(w_1)) \right] ... \left[\prod_{\beta \in A_L} (1 + \cos\beta(w_L)) \right]$$

(for $\gamma \in \Gamma, \cos\gamma = (\gamma + \bar{\gamma})/2$). Observe the following:

$\|\mathbf{R}\|_1 = \hat{\mathbf{R}}(0) = 1$, and $\|\mathbf{R}\|_2 \le \|R\|_\infty \le 2^{|A_1|+\ldots+|A_L|} = 2^{Ls}$.

Hence, for any $1 < p < 2$, by a routine interpolation argument,

$$\|\mathbf{R}\|_p \le \|\mathbf{R}\|_1^{1/p-1/q} \|\mathbf{R}\|_2^{2/q} \le 2^{2.Ls/q} \tag{9}$$

$(1/p + 1/q = 1)$. Also, a routine spectral analysis of R yields

$$\hat{\mathbf{R}}(\beta) = (1/2)^L \tag{10}$$

for all $\beta \in S^L$. Next, let $h = \sum_{\beta \in F \cap (A_1 \times \ldots \times A_L)} \beta,$

whence (by 8)

$$\|h\|_2 = (\Psi_F(s))^{1/2}. \tag{11}$$

Therefore, combining (9), (10) and (11), we obtain

$$\|h\|_2 (\Psi_F(s))^{1/2} 2^{-L} = h * R(0) \le \|h\|_q \|R\|_p \le \|h\|_q 4^{s.L/q}.$$

Letting $q = s$, we obtain the desired inequality.

The following is an immediate consequence of Theorem (5.1.2) and the definition of combinatorial dimension.

Corollary(5.1.3)[226]: Let $F \subset S^L$ be arbitrary. Then. In the case that $\dim F$ is exact,

$$\overline{\lim_{p\to\infty}} \frac{\lambda_F(p)}{p^a} < \infty \tag{12}$$

if and only if $a \ge (\dim F)/2$; in the case that $\dim F$ is asymptotic, (3)holds if and only if

$a > (\dim F)/2$,

Proposition (5.1.4)[226]: Let $F \subset \Gamma$ be arbitrary. The following are equivalent:

(i) $\overline{\lim_{F\to\infty}} \frac{\lambda_F(p)}{p^a} < \infty$;; (ii) for all $f \in L\frac{2}{F}(G)$, $m(|f| > x) < \exp(-Kx^{2/a})$

for all $x > 0$ (m is the Haar measure on $\hat{\Gamma} = G$ and $K > 0$ depends only on F).

Sketch of proof $(\text{i}) \Rightarrow (\text{ii})$ follows by checking that for all $\alpha > 0$

$$\int_G \exp\left(\alpha |f|^{2/a}\right) < \infty. \tag{13}$$

(13) is verified by integrating term by term the Taylor expansion of $\exp\left(\alpha |f|^{2/a}\right)$.

$(\text{ii}) \Rightarrow (\text{i})$ follows by a direct computation of the L^p norm of f.

Here and throughout the section, K (possibly subscripted) will denote a fixed constant whose value may change from one context to another.

Proposition (5.1.5)[226]: Let $F \subset \mathrm{E}^L$ by arbitrary, and suppose that

$$\Psi_F(s) \le Ks^a \tag{14}$$

for all $s > 1$. For every integer $N \ge 1$ and

$$A_1, \ldots, A_L \subset \mathrm{E},\ |A_i| = N, \quad i = 1, \ldots, L,$$

there exists a partition of $F \cap (A_1 \times \ldots \times A_L)$

$$\mathrm{F} = \{F_1, \ldots, F_L\}$$

with the following property: For each $1 \le k \le L$ and all $x \in A_k$

$$\left|\pi_k^{-1}(x) \cap F_k\right| \le KN^{\alpha - 1}. \tag{15}$$

Proof: The proof is by induction on $N \ge 1$. The case $N = 1$ is trivial. Let $N > 1$ and assume the assertion is true for $N - 1$. Let $A_1, \ldots, A_L \subset \mathrm{E}$ be arbitrary, $|A_i| = N, i = 1, \ldots, L$. By (14), we can find $x_i \in A_i, i = 1, \ldots, L$ so that

$$\left|\pi_i^{-1}(x_i) \cap F \cap (A_1 \times \ldots \times A_L)\right| \le KN^{\alpha - 1}.$$

For each i, let $A' = A_i \setminus \{x_i\}$, and apply the induction hypothesis to find a partition $\{F_1', \ldots, F_L'\}$ of $F \cap (A_1' \times \ldots \times A_L')$ so that for all $x \in A_i'$ we have

$$\left|\pi_i^{-1}(x) \cap F_i\right| \le K(N-1)^{\alpha - 1}$$

Let

$$F_i = F_i' \cup \left[\pi_i^{-1}(x_i) \cap F \cap (A_1, \ldots, A_L)\right]$$

for each $i = 1, \ldots, L$. It is easy to verify that $\{F_1, \ldots, F_L\}$ is the required partition.

Corollary (5.1.6)[226]: Let $A_1,...,A_L \subset \mathrm{E}$,

$|A_1| = ... = |A_L| = N > 1$.

Suppose $F \subset A_1 \times ... \times A_L, |F| \geq K_1 N^{\alpha}$ and $\Psi_F(s) \leq K_2 s^{\alpha}$ for all

$s \geq 1$ Then, for some $1 \leq i_0 \leq L$,

$$\left|\left\{x \in A_{i_0} : \left|\pi_{i_0}^{-1}(x)\right| \geq K_1 N^{\alpha-1} / 2L\right\}\right| \geq (K_1 / 4LK_2) N. \qquad (16)$$

In Proposition(5.1.5) we achieved control on cardinality of fibers in F over $x \in A \subset \mathrm{E}$. We now show how to control the cardinality of fibers in $F \subset \mathrm{E}^L$ over $x \in A \subset \mathrm{E}^{L-1}$. In what follows, for each $l \in \{1,...,L\}, \tau_1$ denotes the projection from E^L onto E^{L-1} that is 'orthogonal' to π_1,. The idea for Lemma (5.1.7) below was shown to us by Professor J. Schmerl.

Lemma(5.1.7)[226]: Let $A_1,...,A_L \subset \mathrm{E}$ be arbitrary, $|A_1| = ... = |A_L|$ .Suppose $F \subset A_1 \times ... \times A_L$. Then, there is a partition of F,

$\mathrm{F} = \{F_1,...,F_L\}$

so that for each $l = 1,...,L$ and $x \in \tau_1[A_1 \times ... \times A_L]$,

$$\left|\tau_1^{-1}(x) \cap F_1\right| \leq |F|^{1/L} \qquad (17)$$

Proof: Initialize

$F_1 = ... = F_L = \phi, \quad \mathbb{G} = F$.

Search and sort procedure: Pick a point $x \in \mathbb{G}$ and consider for each $l = 1,...,L$

$$B_1(x) = \tau_1^{-1}(\tau_1(x)) \cap \mathbb{G} \qquad (18)$$

If $|B_l(x)| > |F|^{1/L}$ for each $l = 1,....,L$, place x back in $\mathbb{G}$. Otherwise,

let

$$K = \min\left(l : \left|B_1(x)\right| \leq |F|^{1/L}\right);$$

remove $B_R(x)$ from e and place $B_R(x)$ in F_k.

Repeat this procedure until $B_1(x) > |F|^{1/L}$ for all $l = 1,....,L$ and all $x \in \mathbb{G}$. It is clear that the resulting $F_1,....,F_L$ satisfy (17), and all that is left to prove is the following:

Claim. $\mathbb{G} = \phi$.

Suppose not, and $x \in \mathbb{G}$. From the way the 'search and sort' procedure above is designed, it is clear that

$$\pi_1\left(B_1\left(x\right)\right) \times \ldots \times \pi_L\left(B_L\left(x\right)\right) \subset \mathbb{G}$$

and

$$\left|\pi_1 B_1\left(x\right)\right| > \left|F\right|^{1/L}$$

(recall that π_1 and T_1 are orthogonal projections). But, we then have

$\left|\mathbb{G}\right| > \left(\left|F\right|^{1/L}\right)^L = \left|F\right|$, ,and reach a contradiction.

Proposition (5.1.8)[226]:. Let $F \subset \mathrm{E}^L$ be arbitrary, and suppose that for every $s \geq 1$. $\Psi_F\left(s\right) \leq Ks^a$.

For every integer $N \geq 1$ and $A_1, \ldots, A_L \subset \mathrm{E}, \left|A_i\right| = N, i = 1, \ldots, L,$

there exists a partition of $F \bigcap \left(A_1 \times \ldots \times A_L\right)$, $Z = \left\{F_1, \ldots, F_L\right\}$

with the following property: For each $1 \leq k \leq L$ and all $x \in T_k\left(A_1 \times \ldots \times A_L\right))$

$$\left|T_k^{-1}\left(x\right) \bigcap F_k\right| \leq K_1 N^{(\alpha-1)(L-1)}$$

We recall the definition of a randomly continuous function following ([89] and [37]. An L^2 function on a compact metrizable abelian group $G\left(=\hat{\Gamma}\right)$ $f = \sum_{\gamma \in \Gamma} \hat{f}\left(\gamma\right)\gamma$

is said to be randomly continuous if

$$\left\|f\right\|_{p.s} \equiv \int_{[0,1]} \left\| \sum_{\gamma \in \Gamma} r_\gamma\left(t\right) \hat{f}\left(\gamma\right)\gamma \right\|_{L^\infty(G)} dt < \infty$$

where $\left(r_\gamma\right)_{\gamma \in \Gamma}$ is an enumeration of the usual Rademacher system, i.e. a system of symmetric statistically independent random variables on [0, 1] each of whose range is $\left\{-1, +1\right\}$. The notion of random continuity in the context of harmonic analysis is part of a general philosophy contained in Kahane's monograph, Some Random Series of Functions [127].Next we define the RC-norm of an L-dimensional tensor

$a = \left(a_{i_1 \ldots i_L}\right)_{i_L,,, i_L} \in N$ by

$$\|a\|_{RC}=\sum_{j=1}^{L}\left[\sum_{i_j\in N}\left(\sum_{\substack{i_{k\in N}\\ k\in(1,\dots,L)\\ k\neq j}}\left|a_{i_1},\dots,a_{i_L}\right|^2\right)^{1/2}\right] \tag{19}$$

(Here and throughout, the set of natural numbers denoted by N serves merely as a convenient indexing set.) Returning to $S=\{\beta_n\}_{n\in N}$, the Steinhaus system of independent characters on $\Omega=\otimes T$, and viewing $\hat{f},f\in L_s^2L(\Omega^L)$, as an L-dimensional tensor

$\left(\hat{f}\left(\beta_{i_1},\dots,\beta_{i_L}\right)=a_{i_1\dots i_L}\right)$, our starting point will be the following theorem

Theorem 5.1.9)[226]: For all $f\in L^2s^L\left(\Omega^L\right)$

$$(K_2)^L\left\|\hat{f}\right\|_{RC}\le\|f\|_{p.s}\le(K_1)^L\left\|\hat{f}\right\|_{RC} \tag{20}$$

where $K_1,K_2>0$ are universal constants.

Definition (5.1.10)[226]: The $\rho(q)$-constant of $F\subset\Gamma,2<q<\infty$ is

$\rho_F(q)=\sup\left\{\|f\|_q/\|f\|_{p,s}:f\in L_F^2(G),f\neq 0\right\}$.

Theorem 5.1.11)[226]: Let $F\subset S^L$ be arbitrary. Then: In the case that dim F is exact,

$$\overline{\lim_{q\to\infty}}\frac{\rho_F(q)}{q^a}<\infty \tag{21}$$

if and only if $a\ge(\dim F-1)/2$; in the case that dim F is asymptotic, (21) holds if and only if $a>(\dim F-1)/2$

In order to keep notation as simple as possible, we prove Theorem (5.1.9) in the case $L=2$; the arguments in the general case are similar. In what follows, $F\subset S^2$ will be identified with its underlying indexing set in N^2: $\left\{(n_1,n_2):\left(\beta_{n_1},\beta_{n_2}\right)\in F\right\}\subset N^2$ Slightly abusing notation, we shall occasionally refer to the latter also as F.

Lemma (5.1.12)[226]: Let $F\subset S^2$ and suppose that

$$\Psi_F(s) \le ks^a \qquad (22)$$

for all $s \ge 1$. Then, for all $f \in L_F^2(\Omega^2)$

$$\|f\|_s \le \sqrt{2Ks}^{(a-1)/2} \|\hat{f}\|_{RC} \qquad (23)$$

for all $s \ge 1$.

Proof. Let an arbitrary $f \in L_F^2(\Omega^2)$ be given by

$$f = \sum_{(i_1,i_2)\in F} a_{i_1,i_2}(\beta_{i_1}, \beta_{i_2}).$$

(we identify $F \subset S^2$ with $\{(i_1,i_2) \in N^2 : (\beta_{i_1}, \beta_{i_2}) \in F\} \subset N^2$).. Let $s > 1$ be an arbitrary integer and write

$$f^s = \sum_{(i_{11},i_{21}),\dots,(i_{1s},i_{2s})\in F} a_{i_{11}i_{21}} .. a_{i_{1s}i_{2s}} (\beta_{i_{11}} \dots \beta_{i_{1s}}, \beta_{i_{21}} \dots \beta_{i_{2s}}).$$

We obtain by Plancherel's formula

$$\|f\|_{2s}^{2s} = \sum_{\gamma\in\hat{\Omega}^2} \left| \sum_{\substack{(i_{11},i_{21}),\dots,(i_{1s},i_{2s})\in F \\ (\beta_{i_{11}} \dots \beta_{1s}\beta_{i21} \dots \beta_{i_{2s}})=\gamma}} a_{i_{11}i_{21}} \dots a_{i_{1s}i_{2s}} \right|^2 . \qquad (24)$$

It is clear that the summation in (24) is performed over only those $\gamma's$ which are s-fold products of elements in F. For such $\gamma \in \hat{\Omega}^2$ write $\gamma = ((\beta_{j_{11}} \dots \beta_{j_{1s}}, \beta_{j_{21}} \dots \beta_{j_{2s}}))$ and denote

$C_1(\gamma) = \{j_{11}, \dots, j_{1s}\}$ and $C_2(\gamma) = \{j_{21}, \dots, j_{2s}\}$;

let

$$A(\gamma) = \{(i_1,i_2) \in F : i_1 \in C_1(\gamma) \quad and \quad i_2 \in C(\gamma)\}$$

By the algebraic independence of S, it follows from (24) that

$$\|f\|_{2s}^{2s} \le \sum_{\gamma\in\hat{\Omega}^2} \left| \sum_{\substack{(i_{11},i_{21}),\dots,(i_{1s},i_{2s})\in A(\gamma) \\ (\beta_{i_{11}} \dots \beta_{1s}\beta_{i21} \dots \beta_{i_{2s}})=\gamma}} a_{i_{11}i_{21}} \dots a_{i_{1s}i_{2s}} \right|^2 \qquad (25)$$

By Proposition (5.1.5)and (22), for each $\gamma \in \hat{\Omega}^2$ that participates in the summation in (25), partition $A(\gamma)$ into $A_1(\gamma) = A_1 \subset A(\gamma)$ and $A_2(\gamma) = A_2 \subset A(\gamma)$ so that for each $i_1 \in C_1(\gamma)$ and $i_2 \in C_2(\gamma)$

$$\begin{aligned} \left|\pi_1^{-1}(i_1) \cap A_1(\gamma)\right| &\le Ks^{a-1} \\ \left|\pi_2^{-1}(i_2) \cap A_2(\gamma)\right| &\le Ks^{a-1} \end{aligned} \tag{26}$$

Reassessing (25) in view of the partition above, we obtain

$$\|f\|_{2s}^{2s} \le \sum_{\substack{\varepsilon=(\varepsilon_k)_{k=1}^{s} \\ \varepsilon_k=1,2}} \sum_{\gamma \in \hat{\Omega}^2} \left| \sum_{\substack{(i_{11},i_{21}) \in A_{\varepsilon_1},\ldots,(i_{1s},i_{2s}) \in A_{\varepsilon_s} \\ (\beta_{i_{11}} \ldots \beta_{1s} \beta_{i_{21}},\beta_{i_{2s}}) = \gamma}} a_{i_{11}i_{21}} \ldots a_{i_{1s}i_{2s}} \right|^2 \tag{27}$$

For each $\varepsilon = (\varepsilon_k)_{k=1}^{s}, \varepsilon_k = 1,2$, define a projection T_ε, from $A_{\varepsilon_1}(\gamma) \times \ldots \times A_{\varepsilon_s}(\gamma)$ into N^s by

$T_\varepsilon\left[(i_{11}, i_{21}), \ldots, (i_{1s}, i_{2s})\right] = \left(i_{\varepsilon_1 1}, \ldots i_{\varepsilon s}\right).$

For topographical reasons, we shall write i_{ε_k} for $i_{\varepsilon_k}, k = 1, \ldots, s$, wherever the omission of the second subscript causes no confusion. It follows from (26) that for each $\left(i_{\varepsilon_1}, \ldots, i_{\varepsilon_s}\right)$ in the range of T_ε we have

$$\left|T_\varepsilon^{-1}(i_{11}, i_{21}) \cap A_{\varepsilon_1}(\gamma) \times \ldots \times A_{\varepsilon_s}(\gamma)\right| \le \left(Ks^{a-1}\right)^2 \tag{28}$$

For each $\varepsilon = (\varepsilon_k)_{k=1}^{s}$, define $\delta_\varepsilon = \delta(\delta_k)_{k=1}^{s}$ by

$$\delta_k = \begin{cases} 1 if\, \varepsilon_k = 2 \\ 2 if\, \varepsilon_k = 1 \end{cases}$$

$$\sum_{\gamma \in \Omega^2} \left| \sum_{\substack{((i_{11},i_{21}),\ldots,(i_{15},i_{25})) \in A_{\varepsilon_1} \times \ldots \times A_{\varepsilon_s} \\ (\beta_{i_{11}} \ldots \beta_{i_{15}}, \beta_{i_{21}} \ldots \beta_{i_{25}}) = \gamma}} a_{i_{11}i_{21}} \ldots a_{i_{15}i_{25}} \right|^2$$

For a fixed $\varepsilon = (\varepsilon_k)_{k=1}^{s}$, write

$$= \sum_{\gamma \in \Omega^2} \left| \sum_{\substack{i_{\varepsilon_1},\ldots,i_{\varepsilon_s} \\ (\beta_{i_{11}} \ldots \beta_{1s} \beta_{i_{21}},\beta_{i_{2s}}) = \gamma}} \left(\sum_{i_{\varepsilon_1}, i_{\delta_s}} a_{i_{11}i_{21}} \ldots a_{i_{1s}i_{2s}} \right) \right|^2 \tag{29}$$

(we write i_{δ_k} for $i_{\delta_k k}$)

Applying (28) and Schwartz's inequality to the third summation (over $i_{\delta_1},...,i_{\delta_s}$) in (29), we obtain that (29) is majorized by

$$Ks^{(a-1)^s}\left|\sum_{\gamma\in\Omega^2}\sum_{\substack{i_{\varepsilon_1},i_{\varepsilon s}\\ \left(\beta_{i_{11}}...\beta_{1s}\beta_{i_{21}},\beta_{i_{2s}}\right)=\gamma}}\left(\sum_{i_{\delta_1},i_{\delta s}}\left|a_{i_{11}i_{21}}...a_{i_{1s}i_{2s}}\right|^2\right)^{1/2}\right|^2 \quad (30)$$

Claim. (30) is majorized by

$$\left[Ks^{(a-1)}\right]^s\left|\sum_{i_{\varepsilon_1}...i_{\varepsilon s}}\left(\sum_{i_{\delta_1},...,i_{\delta_s}}\left|a_{i_{11}i_{21}}...a_{i_{1s}i_{2s}}\right|^2\right)\right|^2 \quad (31)$$

which, in turn, equals

$$\left[Ks^{(a-1)}\right]^s\left\{\left[\sum_{i_{\varepsilon_1}}\left(\sum_{i_{\delta_1}}\left|a_{i_1i_2}\right|^2\right)^{1/2}\right]...\left[\sum_{i_{\varepsilon s}}\left(\sum_{i_{\delta_s}}\left|a_{i_1i_2}\right|^2\right)^{1/2}\right]\right\}^s \quad (32)$$

(The summations in (31) and (32) are performed freely over $i_{\varepsilon_1},...,i_{\varepsilon_s}\in N,$ and $i_{\delta_1},...,i_{\delta_s}\in N$ respectively.)

To establish the claim, we first note that each $\gamma\in\hat{\Omega}^2$ that is a product of s elements in S X S can be viewed as

$\gamma=\gamma_\varepsilon.\gamma_\delta,$

where γ_ε and γ_δ' are products of s elements in $(\{\beta_0\}\times S)\cup(S\times\{\beta_0\})$ respecting the following scheme:

$$(a)\begin{cases}\gamma_\varepsilon = \gamma_1...\gamma_s\\ \gamma_\delta = \gamma_1'...\gamma_s'\end{cases}$$

$$(b)\begin{cases}\gamma_k = (\gamma_k(1),\gamma_k(2))\in(\{\beta_0\}\times E)\cup(E\times\{\beta_0\})\\ \gamma_k' = (\gamma_k'(1),\gamma_k'(2))\in(\{\beta_0\}\times E)\cup(E\times\{\beta_0\})\end{cases}$$

where the δ_k^{th} coordinate of 'γ_k is β_0 and the ε_k^{th} coordinate of γ_k' is $\beta_0, K=1,...,s.$,.Let

$A(\gamma_\varepsilon)=\{(i_{\varepsilon_1},...,i_{\varepsilon_s}):\gamma_\varepsilon=\gamma_1...\gamma_s$.

as in (a) and (b), and $\gamma_k(\gamma_k)=\beta_{i_{\varepsilon_k}}, k=1,...,s\}$ Similarly, $A(\gamma_\delta)=\{(i_{\delta_1},...,i_{\delta_s}):\gamma_\delta=\gamma'_1...\gamma'_s$ as in (a) and (b), and $\gamma'_k(\delta_k)=\beta_{i_{\delta_k}}, k=1,...,s\}$. Next, observe that (30) is majorized by

$$\left(Ks^{(as-1)}\right)^s \sum_{\gamma_\varepsilon}\sum_{\gamma_\delta}[\sum_{A(\gamma_\varepsilon)}(\sum_{A(\gamma_\delta)}\left|a_{i_{11}i_{21}}...a_{i_{1s}i_{2s}}\right|^2)^{1/2}]^2 \tag{33}$$

($\sum_{\gamma_\varepsilon} and \sum_{\gamma_\delta}$, are summations over $\gamma_\varepsilon,\gamma_\delta\in\hat{\Omega}^2$ described by (a) and (b) above; $\sum_{A(\gamma_\varepsilon)} and \sum_{A(\gamma_\delta)}$ are summations over $\left(i_{\varepsilon_1},...i_{\varepsilon_s}\right)$ and $\left(i_{\delta_1},...,i_{\delta_s}\right)$ taking values in $A(\gamma_\varepsilon)$ and $A(\gamma_\delta)$, respectively.) (33) is certainly majorize d by

$$\left(Ks^{(a-1)}\right)^s \left|\sum_{\gamma_\varepsilon}\{\sum_{\gamma_\delta}[\sum_{A(\gamma_\varepsilon)}(\sum_{A(\gamma_\delta)}|.|^2)^{1/2}]^2\}^{1/2}\right|^2,$$

which, by an application of Minkowski's inequality, is majorized by

$$\left(Ks^{(a-1)}\right)^s \left|\sum_{\gamma_\varepsilon}\sum_{A(\gamma_\varepsilon)}\{\sum_{\gamma_\delta}\sum_{A(\gamma_\delta)}|.|^2\}^{1/2}\right]^2. \tag{34}$$

Clearly, $A(\gamma)\cap A(\gamma'_\varepsilon)=A(\gamma_\delta)\cap A(\gamma'_\delta)=\phi$ whenever $\gamma_\varepsilon\neq\gamma'_\varepsilon$ and $\gamma_\delta\neq\gamma'_\delta$ and therefore (34) equals

$$\left(Ks^{(a-1)}\right)^s\{\sum_{i_{\varepsilon_1},...,i_{\varepsilon_s}}(\sum_{i_{\delta_1},...,i_{\delta_s}}\left|a_{i_{11},i_{21}}...a_{i_{1s},i_{2s}}\right|^2)^{1/2}\}^2$$

and that completes the proof of the claim. Each of the s factors in (32) is majorized by $\left\|\hat{f}\right\|_{RC}$, and we obtain from the claim and (27) that

$$\left\|\hat{f}\right\|_{2s}\leq 2^s\left(Ks^{(a-1)}\right)^s\left\|\hat{f}\right\|_{RC}^{2s}$$

which completes the proof of (23).

Lemma 5.1.13)[226]: Let $F\subset S^2$ and suppose that

$$\overline{\lim_{s\to\infty}}\frac{\Psi_F(s)}{s^\eta}=\infty. \tag{35}$$

Then: For $D>0$ and integers $s>1$ as large as we please there are $h\in L_F^2\left(\Omega^2\right)$ so that

$$\left(s^{(\eta-1)/2}\left\|\hat{h}\right\|_{RC}\right)D^{1/2}\leq\|h\|_s. \tag{36}$$

Proof: Let D be as large a number as we please. By (35) we can find s > 1 as large an integer as we please, and $A, B \subset S, |A| = |B| = s$ so that

$F \bigcap (A \times B) \geq Ds^{\eta}$.

Without loss of generality, assume that $V \equiv F \bigcap (A \times B)$ contains $\left[Ds^{\eta}\right]$ points ($\left[Ds^{\eta}\right]$ denotes here the largest integer smaller than Ds^{η}). Let

$$h = \sum_{(\beta_1, \beta_2) \in V} (\beta_1, \beta_2).$$

We clearly have

$$\|h\|_2 \leq 2D^{1/2} s^{\eta/2} \tag{37}$$

which implies the following

Claim.

$$\left\|\hat{h}\right\|_{RC} \leq KD^{1/2} s^{(\eta+1)/2}$$

(as usual, K denotes a fixed constant).

Proof of claim: It follows from (37)and the Kahane-Salem Zygmund probabilistic estimates of the sup-norm of random trigonometric polynomials that there is a choice of signs $\pm$ for which

$$\left\| \sum_{(\beta_1, \beta_2) \in V} \pm (\beta_1, \beta_2) \right\|_{\infty} \leq KD^{1/2} s^{\eta/2} s^{1/2} \tag{38}$$

Denote the characteristic function of V by χ_V and obtain the left hand side

$$\text{of } (38) = \sup_{\omega \in \Omega} \sum_{\beta_1 \in A} \left| \sum_{\beta_2 \in B} XV(\beta_1, \beta_2) \beta_2(\omega) \right|$$

$$\geq \int_{\Omega} \sum_{\beta_1 \in A} \left| \sum_{\beta_2 \in B} XV(\beta_1, \beta_2) \beta_2(w) \right| dw$$

$$\geq (1/C) \sum_{\beta_1 \in A} [\sum_{\beta_2 \in B} XV(\beta_1, \beta_2)]^{1/2}$$

The last inequality above was obtained by an application of the Khintchin inequality for the Steinhaus system (C above is the Khintchin constant of S whose precise determination is still an open problem). The roles of β_1 and β_2

are interchangeable in the estimation above, and the claim is thus established. Let $\mathbf{R}$ be the Riesz product

$$\mathbf{R}(w_1,w_2)=\left[\prod_{\beta_1\in A}(1+\cos\beta_1(w_1))\right],\left[\Pi_{\beta_1\in B}\left(1+\cos\beta_2\left(w_2\right)\right)\right]$$

and as in the proof of Theorem 5.1.2)[226]: we conclude

$$\left(\frac{D}{4}\right)s^{\eta}\le R*h(0)\le 8\|h\|_s \tag{39}$$

Combining the claim and (39) we obtain (36).

Combining Lemma (5.1.12) Lemma (5.1.13)and Theorem (5.1.9) we obtain Theorem (5.1.11).

The application of the decomposition property given by Proposition (5.1.5) is a crucial step in the proof of Theorem (5.1.11) (see (26)above).Following the line of arguments that is completely analogous to the proof of Lemma (5.1.13) via the decomposition property given by Proposition (5.1.8) we obtain Lemma(5.1.14) below. First, some notation: Let $a=\left(a_{n_1,\ldots,n_L}\right)_{n_1,\ldots,n_L}\in N$ be an L-dimensional tensor. Define the $R\tilde{C}$-norm of a by

$$\|a\|_{R\bar{C}}=\sum_{j=1}^{L}[\sum_{\substack{i_k\in N\\ k\in(1,\ldots,L)\\ k\neq j}}(\sum_{i_j\in N}\left|a_{i_1\ldots i_L}\right|^2)^{1/2}]$$

Lemma(5.1.14)[226]: Let $F\in S^L$ be so that $\Psi_F(s)\le K_1 s^{\alpha}$ for all $s\ge 1$. Then, for every $f\in L_F^2\left(\Omega^L\right)$

$$\|f\|_q\le K_2 q^{(a-1)/2(L-1)}\left\|\hat{f}\right\|_{R\bar{C}}$$

for all $q>2$.

We start by recalling the classical Littlewood and Orlicz inequalities whose statements given here are slightly different from the ones given in [112] and [266]: For all f continuous functions on Ω^2 with spectrum in $S^2\left(=Cs^2\left(\Omega^2\right)\right)$

$$\sum_{\beta_1\in s}(\sum_{\beta_2\in s}\left|\hat{f}\left(\beta_1,\beta_2\right)\right|^p)^{1/p}<\infty,\ \text{if and only if } p\ge 2,\ \text{(Littlewood)} \tag{40}$$

$$\sum_{\beta_1\in s}(\sum_{\beta_2\in s}\left|\hat{f}(\beta_1,\beta_2)\right|)^p<\infty \text{ if and only if } p\geq 2. \qquad \text{Orlicz} \qquad (41)$$

$$\sum_{\beta_1,\beta_2\in S^2}\left|\hat{f}(\beta_1,\beta_2)\right|^p<\infty \text{ if and only if } p\geq 4/3. \qquad \text{(Littlewood)} \qquad (42)$$

Still within a classical context, the following are multidimensional extensions of the statements above: Let $L>1$ be an arbitrary integer. For all

$f\in Cs^L\left(\Omega^2\right)$

$$\sum_{\beta_1\in S}(\sum_{(\beta_2,\ldots,\beta_L)\in s^{L-1}}\left|\hat{f}(\beta_1,\ldots,\beta_L)\right|^p)^{1/p}<\infty \text{ if and only if } p>2; \qquad \text{(i)}$$

$$\sum_{(\beta_2,\ldots,\beta_L)\in s^{L-1}}(\sum_{\beta_1\in S}\left|\hat{f}(\beta_1,\ldots,\beta_L)\right|)^p \text{ if and only if } p>2; \qquad \text{(ii)}$$

$$\left\|\hat{f}\right\|_p\equiv(\sum_{(\beta_1,\ldots,\beta_L)\in S^L}\left|\hat{f}(\beta_1,\ldots,\beta_L)\right|^p)^{1/p} \text{ if and only if } p\geq 2/\left(1+\frac{1}{L}\right). \qquad \text{(iii)}$$

(i) and (ii), straightforward extensions of (40) and (41), appear in the literature on ad hoc basis; $(\text{ii})_L$ was obtained in [90]. In this section, we establish 'continuous' systems of inequalities in which (i) and (ii) are 'discrete' instances. First some notation: In what follows, we shall consider norms of restrictions of $\hat{f}$ to $F\subset S^L$ denoted by $\hat{f}X_F$. For example,

$$\left\|\hat{f}X_F\right\|_s\equiv(\sum_F\left|\hat{f}\right|^2)^{1/2}=(\sum_{(\beta_1,\ldots,\beta_L)\in S^L}\left|\hat{f}(\beta_1,\ldots,\beta_L)X_F(\beta_1,\ldots,\beta_L)\right|^s)^{1/2}$$

Where X_F is the characteristic function of F.

Theorem (5.1.15)[226]: (An extension of Orlicz's inequality). Let $F\subset S^2$ be arbitrary, dim F exact (respectively, dim F asymptotic). Then: For all $f\in Cs^2\left(\Omega^2\right)$

$$\sum_{\beta_1\in S}(\sum_{\beta_2\in S}\left|\hat{f}(\beta_1,\beta_2)X_F(\beta_1,\beta_2)\right|)^p<\infty$$

if and only if

$$p\geq 2/(3-\dim F)\left(\textit{respectively}, p>2/(3-\dim F)\right).$$

Theorem (5.1.16)[226]: (an extension of Littlewood's inequality, (42) above. Let $F\subset S^2$ be arbitrary, dim F exact (respectively, dim F asymptotic). Then: for all

$f \in Cs^L\left(\Omega^L\right)$

$$\sum_{\beta_1,\ldots,\beta_L \in S}\left|\hat{f}\left(\beta_1,\ldots,\beta_L\right) X_F\left(\beta_1,\ldots,\beta_L\right)\right|^p < 0$$

if and only if

$$p \geq 2 \Big/ \left(1+\frac{1}{\dim F}\right)\left(\text{respectively,} p > 2 \Big/ \left(1+\frac{1}{\dim F}\right)\right)$$

The proofs make use of the results the Kahane-Salem-Zygmund estimates and an in-stance of a general theorem due to in[89])). To establish that $p > 2/(3-\dim F)$ and $p \geq 2/(1+1/\dim F)$ are sufficient in Theorems (5.1.15) and (5.1.16) respectively (with strict inequality in the asymptotic dim F case), we follow the strategy of the proof) in [89].

Lemma(5.1.17)[226]: Let $F \subset S^2$ and suppose $\Psi_F(s) \leq Ks^{\alpha}$ for all $. s > 0$

(i) Let $\phi \in l^{\infty}\left(S^2\right), \phi \equiv 0 \, on S^2 \setminus F,$, and

$$\sum_{\beta_1 \in S}\left(\sup_{\beta_2 \in S}\left|\phi\left(\beta_i, \beta_j\right)\right|\right)^{2(a-1)} \leq 1,$$

$i, j = 1$ and $i \neq j$ then,

$$\overline{\lim_{p \to \infty}} \frac{\left\|\sum_{\beta \in F} \phi(\beta) \hat{f}(\beta) \beta\right\|_p}{p^{1/2}} < \infty$$

for all $f \in L_F^2\left(\Omega^2\right)$.

(ii) Let $\phi \in l^{2/(a-1)}\left(S^L\right), \phi \equiv 0 \, on \, S^L \setminus F.$ Then ,

$$\overline{\lim_{p \to \infty}} \frac{\left\|\sum_{\beta \in F} \phi(\beta) \hat{f}(\beta) \beta\right\|_p}{p^{1/2}} < \infty$$

for all $f \in L_F^2\left(\Omega^2\right)$. .

Proof: (i) Let $f \in L_F^2\left(\Omega^2\right)$. and fix an arbitrary $p > 2$. Define

$$\phi_1 = \begin{cases} \phi & if\ |\phi| \le p^{(1-\alpha)/2} \\ 0 & otherwise, \end{cases}$$

and

$\phi_2 = \phi - \phi_1$.

A straightforward computation yields

$$\sum_{\beta_1 \in S} s \left(\sup_{\beta_{2 \in S}} \right) \left| \phi_2 \left(\beta_1, \beta_2 \right)^2 \right|^{1/2} \le p^{(2-\alpha)/2.} \tag{43}$$

We estimate

$$\left\| \sum_{\beta \in F} \phi(\beta) \hat{f}(\beta) \beta \right\|_p \le \left\| \sum_{\beta \in F} \phi_1(\beta) \hat{f}(\beta) \beta \right\|_p + \left\| \sum_{\beta \in F} \phi_2(\beta) \hat{f}(\beta) \beta \right\|_p$$

By the assumption on F and Theorem(5.1.2) we obtain

$$\left\| \sum_{\beta \in F} \phi_1(\beta) \hat{f}(\beta) \beta \right\|_p \le \left(K p^{\alpha} \right)^{1/2} \left\| \sum_{\beta \in F} \phi_1(\beta) \hat{f}(\beta) \beta \right\|_2$$

$$\le \sqrt{K} p^{1/2} \|f\|_2 . \tag{44}$$

From (43) we deduce

$$\left\| \phi_2 \hat{f} \right\|_{RC} \le p^{(2-\alpha)/2} \|f\|_2$$

and thus by Lemma(5.1.13) we have

$$\left\| \sum_{\beta \in F} \phi_2(\beta) \hat{f}(\beta) \beta \right\|_p \le K p^{\alpha 1/2} \|f\|_2 \tag{45}$$

The conclusion in (1) follows from (44) and (45).The proof of (2) is practically identical and will be omitted.

are necessary in Theorems(5.1.15) and(5.1.16) (strict inequality when dimF asymptotic): Suppose there exists C > 0 so that for all $f \in C_F\left(\Omega^2\right)$.

$$C \|f\|_\infty \ge \left(\sum_{\beta_1 \in S} \left(\sum_{\beta_1 \in S} \left| \hat{f}\left(\beta_1, \beta_2\right) \right| \right)^p \right)^{1/p} . \tag{46}$$

Write dim $F = \alpha$. Suppose $s > 1, A_1, A_2 \subset S, |A_1| = |A_2| = s$ and $\left| \left(A_1 \times A_2 \right) \cap F \right| = K s^{\alpha}$.

By Corollary (5.1.6) we assume without loss of generality that for all $\beta_1 \in A_1$

$$\left|\pi_1^{-1}(\beta_1)\right| \geq Ks^{\alpha-1}. \quad (47)$$

Write $A = (A_1, A_2) \cap F$.. We obtain a choice of signs $\pm$ so that

$$\left\| \sum_{(\beta_1,\beta_2)\in A} \pm(\beta_1, \beta_2) \right\|_2 \leq K_1 s^{(\alpha+1)/2} \quad (48)$$

$K_1 > 0$ depends only on K). Combining (48), (47) and (46), we deduce

$$K_1 Cs^{(\alpha+1)/2} \geq \left(\sum_{\beta_1 \in A} \left(\sum_{\beta_1 \in A_2} \times A(\beta_1, \beta_2) \times F(\beta_1, \beta_2) \right)^p \right)^{1/p} \quad (49)$$

$$\geq Ks^{1/p} s^{p(\alpha-1)}.$$

(49) holds for arbitrarily large s only if $p \geq 2/(3-\alpha)$.If $\lim_{s\to\infty} \left(\Psi_F(s)/s^{\alpha} \right) = \infty$

then (49) implies that $p > 2/(3-\alpha)$. This completes the proof of Theorem (5.1.15) The proof that $p \geq 2/\left(1+\left(1/\dim F\right)\right)$ is necessary in (5.1.16) Theorem is similar and will be omitted.

Theorem (5.1.18)[226]: (An extension of Orlicz's inequality). Let $F \subset S^L$ be arbitrary, dim F exact (respectively, dim F asymptotic). Then: For all $f \in Cs^L(\Omega^L)$

$$\sum_{\beta_1,\ldots,\beta_{L-1}\in s} \left(\sum_{\beta_L \in s} \left| \hat{f}(\beta_1,\ldots,\beta_2) X_F \right|\right)^p < \infty$$

for all

$$p \geq 2((L-2)\dim F + 1) / ((L-2)\dim F + L - \dim F + 1)$$

(in the asymptotic dim F case, the above is a strict inequality).Depending on the 'combinatorial' structure of $F \subset S^L$, Theorem (5.1.18) may or may not be sharp. We illustrate:

(i) If dim $F = L$, then the inequality in Theorem (5.1.18) reduces to the usual Orlicz inequality (41)L which is sharp.

(ii) If $L = 2$, then the inequality in Theorem (5.1.18) reduces to the one in Theorem (5.1.15) which is sharp.

Theorem (5.1.19)[226]: (Another extension of Orlicz's inequality).Suppose $F = S^J \times F_1, J \ge 1$ and $F_1 \subset S^2$ Then: For all $f \in C_F\left(\Omega^{J+2}\right)$

$$\sum_{(\beta_1,\dots,\beta_{J+1-1})\in S^{J+1}} \Big(\sum_{\beta_{J+2}\in S} \left|\hat{f}(\beta_1,\dots,\beta_{J+2})\right|^p \Big)^{1/p} < \infty$$

if and only if(i) $p \ge (2J+2)/(2J+3-\dim F)$, dim F exact, (ii) $p > 2J+2/(2J+3-\dim F)$, dim F asymptotic).

The proof of Theorem (5.1.19) follows the line of arguments used in the proof of

Theorem (5.1.15) In the case $F = S^J \times F_1, J \ge 1$ 1and $1 < \dim F_1 < 2$, the inequality in Theorem (5.1.19) is sharper than the one given by Theorem (5.1.18) The classical Orlicz inequality (for all $f \in Cs^2\left(\Omega^2\right)$,

$$\sum_{\beta_1\in s} \Big(\sum_{\beta_2\in s} \left|\hat{f}\left(\beta_1,\beta_2\right)\right| \Big)^2 < \infty$$

follows from the classical Littlewood inequality (for all $f \in Cs^2\left(\Omega^2\right)$,

$$\sum_{\beta_1\in s} \Big(\sum_{\beta_2\in s} \left|\hat{f}\left(\beta_1,\beta_2\right)\right|^2 \Big)^{1/2} < \infty)$$

which, in turn, follows from the classical Khintchin inequality (for all

$$f \in L_s^1\left(\Omega\right), \|f\|_1 \ge K \|f\|_2);$$

in fact, these three inequalities are 'equivalent'(see [230]). The extended Orlicz inequality of Theorem (5.1.15) however, does not follow from or imply a Littlewood-type inequality:

Proposition (5.1.20)[226]: Suppose $F \subset S^2, \dim F > 1$.For every $1 \le p < 2$ there is $f \in C_F\left(S^2\right)$ So that

$$\sum_{\beta_1\in s} \Big(\sum_{\beta_2\in s} \left|\hat{f}\left(\beta_1,\beta_2\right)\right|^p \Big)^{1/p} = \infty.$$

Proof: We argue as we did to establish the necessity of $p \ge 2/(3-\dim F)$

in Theorem (5.1.16) Suppose that $s > 1, A_1, A_2 \subset S |A_1| = |A_2| = s$ and

$$\left|\left(A_1 \times A_2\right) \cap F\right| = Ks^\alpha$$

where $\alpha > 1$. We follow (47) and (48) but in place of (49) we write

$$K_1 C s^{(\alpha+1)/2} \geq \sum_{\beta_1 \in A} \left(\sum_{\beta_2 \in A_2} X_A(\beta_1, \beta_2) \right)^{1/p} \geq K s\, s^{(\alpha-1)/p}.$$

The inequality above is valid for arbitrarily large s only if $p \geq 2$.

We now move to a general harmonic analytic setting. With the aim of simplifying future arguments, we start by altering slightly the definition of $\Psi_F F \subset \mathrm{E}^L$, (given in (1)): Let $s \geq 1$ and write

$$\Psi_F(s) = \max\left\{ \left| F \cap A^L \right| : A \subset \mathrm{E}, |A| = s \right\}. \tag{50}$$

It is trivial to see that the redefinition of Ψ_F has no impact on the definition of $\dim F$

(Definition (5.1.1).We recall: $E \subset \Gamma$ is said to be K-independent, K a positive integer, if for any $J \geq J' > 0$, and $\{\gamma_1, ..., \gamma_J\} \subset E$,, the relation

$$\prod_{j=1}^{J} \gamma_j^{\lambda_j} = \prod_{j=1}^{J'} \gamma_j^{v_j}$$

where the Xi's and the vj's are integers in $[-K, K]$, implies that $J = J'$ and $\lambda_j = V_j$ for $j = 1, ..., J.E$ is independent if it is K-independent for every K; 1-independent sets are traditionally called dissociate sets. From here on, $\mathrm{E} \subset \Gamma$ will denote an infinite dissociate set which does not contain 1_Γ,the identity element of Γ. Fix an arbitrary integer $L > 0$, and define $\mathrm{E}_L = \{\gamma_1^{\varepsilon 1} ... \gamma_L^{\varepsilon L} : \gamma_1, ..., \gamma_L$ distinct characters in E}.More generally, fixing $\varepsilon = \left(\varepsilon_j\right)_j^L = 1$, $\varepsilon_j = \pm 1$, we define distinct characters in E }.Finally, define

$$|\mathrm{E}|_L = \left(\bigcup_{k=1}^{L} \bigcup_{\substack{\varepsilon = (\varepsilon_j)_{j=1}^{k} \\ \varepsilon_j = \pm 1}} \mathrm{E}_k^{\varepsilon} \right) \cup \{1_\Gamma\}$$

Next, we identify $[\mathrm{EI}_\mathrm{L}$ with a subset of the L-fold Cartesian product, $(\mathrm{E} \cup \mathrm{E}^{-1} \cup \{1_\Gamma\})^L : \gamma = \gamma_1^{\varepsilon_1} ... \gamma_k^{\varepsilon_k} [\mathrm{E}]_L$ is identified with $\overline{\gamma} = \left(\gamma_1^{\varepsilon_1}, .., \gamma_k^{\varepsilon_k}, 1_\Gamma, ..., 1_\Gamma\right) \in \Gamma^L$ We designate

$$\left[\tilde{\mathrm{E}}\right]_L = \left\{ \tilde{\gamma} : \gamma \in [\mathrm{E}]_L \right\} \subset \Gamma^L;$$

given $F \subset [\mathrm{E}]_L \subset \Gamma$,, we denote

$$\tilde{F} = \{\tilde{\gamma} : \gamma \in F\} \subset \left(\mathrm{E} \cup \mathrm{E}^{-1} \cup \{1_\Gamma\}\right)^L \qquad (51)$$

Definition (5.1.21)[226]: Let $.F \subset \Gamma$

(i) The A-exponent of F is given by

$$\theta_F = \inf\left\{a : \overline{\lim_{p\to\infty}} \frac{\lambda_F(p)}{p^a} < \infty\right\}.$$

θ_F is exact if $\overline{\lim_{p\to\infty}}\left(\lambda_F(p)/p^{\theta_F}\right) < \infty$; ; otherwise, θ_F is asymptotic.

(ii) The p-exponent of F is given by

$$r_F = \inf\left\{a : \overline{\lim_{q\to\infty}} \frac{\rho_F(q)}{q^a} < 0\right\}.$$

r_F is exact if $\overline{\lim}_{q\to\infty}\left(\rho_F(q)/q^{r_F}\right) < 0$; otherwise, r_F is asymptotic. (The definition of $\lambda_F(p)$,

the $A(p)$ constant of F, is stated in Corollary (5.1.3) the definition of $\rho(q)$ is given in Definition (5.1.10)

Definition (5.1.22)[226]: $F \subset \Gamma$ is a p-Sidon set (respectively, asymptotic p - Sidon set) if

$$C_F\left(\hat{\Gamma}\right) \subset L^\gamma \qquad (52)$$

if and only if $r \ge p$ (respectively, $r > p$). (Following tradition, we refer to 1-Sidon sets as Sidon sets.) The Sidon exponent of F is given by $\sigma_F = p$ and is exact if F is p-Sidon, and asymptotic if F is asymptotically p-Sidon. Recall that the Sidon constant of $F \subset \Gamma$ is given by

$$K_F = \sup\left\{\left\|\hat{f}\right\|_1 / \|f\|_\infty : f \in C_F\left(\hat{\Gamma}\right), f \ne 0\right\}$$

For each positive integer n, define

$$\Phi_F(n) = \sup\{K_A : A \subset F, |A| = n\}.$$

Definition (5.1.23)[226]: The Sidon characteristic of $F \subset \Gamma$ is given by

$$\eta_F = \inf\left\{a : \overline{\lim_{n\to\infty}} \frac{\Phi_F(n)}{n^a} < \infty\right\}.$$

η_F is exact if $\overline{\lim}_{n\to\infty}\left(\Phi_F(n)/n^{\eta_F}\right)$; otherwise, η_F is asymptotic.

It is easy to see that $0 \le \eta_F \le 1/2$ and we note two obvious external cases: (i) $\eta_F = 0$ is exact if and only if F is 1-Sidon. (ii) $\eta_\Gamma = \frac{1}{2}$ is exact.

The first statement is a trivial tautology. The second statement, appropriately translated, is folklore in various contexts (e.g., see section 1.6 in [24]).

Lemma (5.1.24)[226]: Let $F \subset \mathrm{E}_L$. For every integer s > 1

$$\lambda_F(2s) \le 2^L \left(\Psi_F(sL)\right)^{1/2}.$$

Proof: We shall prove the lemma first in the particular case where E is an infinite independent set in some Γ. For example, we can take E to be the Steinhaus system. For the purpose of the proof, designate

$$\bar{F} = \left\{(\beta_1,...,\beta_L) \in \mathrm{E}^L : \beta_1...\beta_L \in F\right\} \tag{53}$$

Clearly, $\bar{F} \subset \mathrm{E}^L$ is symmetric:

$$(\beta_1,...,\beta_L) \in \bar{F} \Rightarrow \left(\beta_{r(L)},...,\beta_{r(1)}\right) \in \bar{F}$$

for any τ, a permutation of $\{1,...,L\}$. We thus trivially have

$$L!\Psi_{\tilde{F}} = \Psi_{\bar{F}} \tag{54}$$

($\Psi_{\tilde{F}}$ and $\Psi_{\bar{F}}$ are given by (50)). Let $f \in L^2_F(G)$, $f = \sum_{\beta\in F}\hat{f}(\beta)\beta$, which can be rewritten as

$$f = \frac{1}{L!}\sum_{\beta\in\bar{F}} a_\beta \pi_1(\beta)...\pi_L(\beta)$$

Where $a_\beta = \hat{f}\left(\pi_1(\beta)...\pi_L(\beta)\right)$ (as usual $\pi_1,...,\pi_L$ denote the canonical projections from E^L into E). As in the proof of Theorem (5.1.2) let s > 1 be an arbitrary integer and write

$$\|f\|_{2s}^{2s} = \left(\frac{1}{L!}\right)^{2s} \sum_{\gamma\in\Gamma}\left|\sum_{A(\gamma)} a_{\beta_1}...a_{\beta_s}\right|^2, \tag{55}$$

Where $A(\gamma) = \left\{(\beta_1,...,\beta_s) \in \bar{F} : \pi_1(\beta_1)...\pi_L(\beta_1)...\pi_L(\beta_s) = \gamma\right\}$.

Following an argument similar to the one in the proof of Theorem (5.1.2) we deduce that

$A(\gamma)\le(\Psi_{\bar{F}}(sL))^{s}$ and thus obtain from (55)

$$\|f\|_{2s}\le(\Psi_{\bar{F}}(sL))^{1/2}\left(\frac{1}{L!}\right)(\sum_{\gamma\in\bar{F}}|a_\beta|^2)^{1/2},$$

But,

$$\|f\|_{2}=\left(\frac{1}{L!}\right)^{1/2}(\sum_{\gamma\in\bar{F}}|a_\beta|^2)^{1/2},$$

and we obtain from (54)

$$\|f\|_{2s}\le(\Psi_{\bar{F}}(sL))^{1/2}\|f\|_2 \quad (56)$$

To prove the lemma in the case $F\subset \mathrm{E}_L$, dissociate $\mathrm{E}=\{\gamma_n\}_{n\in N}\}$ we employ a Riesz product argument and make a reduction to the independent case considered above (see [228], for example). Let $\mathrm{E}_0=\{\beta_n\}_{n\in N}\subset\Gamma_0\left(=\hat{G}_0\right)$ be an infinite independent set of characters. For $\gamma=\gamma_n\ldots\gamma_{n_L}\in F$ denote $\beta_\gamma=\beta_{n_1}\ldots\beta_{n_L}$ and $F_0=\{\beta_\gamma:\gamma\in F\}\subset(\mathrm{E}_0)$ Let $f\in L^2_F(G)$ be arbitrary, $f=\sum_{\gamma\in F}\hat{f}(\gamma)\gamma$. Fix $t\in G_0$ and let $f_t=\sum_{\gamma\in F}\hat{f}(\gamma)\overline{\beta_\gamma(t)}\gamma$,

where $\overline{\quad}$ denotes here complex conjugation. Next, write a Riesz product

$$\mu_t=\prod_{n=1}^{\infty}\left(1+\frac{\beta_n(t)\gamma_n+\overline{\beta_n(t)}\overline{\gamma}_n}{2}\right)$$

We easily have $\|\mu_t\|_M=1$ and $u_t*f_t=f/2^L$, and thus obtain from (56)

$$\|f\|_{2s}\le 2^L(\Psi_{\bar{F}}(sL))^{1/2}\|f\|_2 \text{ for all } s>1..$$

Lemma (5.1.25)[226]: Let $F\subset\mathrm{E}_L$. Then, for every integer $s>1$.

$$2^{-L-1}(\Psi_F(s))^{1/2}\le\lambda_F(2s).$$

Proof: Let $A\subset\mathrm{E},|A|=s$ be so that $\Psi_F(s)=|A^L\cap\tilde{F}|$, and define $f=\sum_{\gamma\in A^L\cap\tilde{F}}\pi_1(\gamma)\ldots\pi_L(\gamma)$, whence

$$\|f\|_2=(\Psi_{\bar{F}}(s))^{1/2}$$

Define the Riesz product $\mathbf{R} = \prod_{\gamma \in A} (1 + \cos\gamma)$

$(\cos\gamma = (\gamma + \bar{\gamma})/2)$ whence $\|\mathbf{R}\|_1 = 1$, $\|\mathbf{R}\|_2 \le 2^s$ and $\|\mathbf{R}\|_p \le 2^{2s/q}$, $1/p + 1/q = 1$, for all $1 < p < 2$. We therefore have

$$2^{-L} \Psi_{\bar{F}}(s) = \mathbf{R} * f(0) \le 2^{2s/q} \|f\|_q .$$

Setting $q = s$ in the inequality above, we obtain the lemma.

Lemma (5.1.26)[226]:

$$\|\tilde{f}\|_\infty \le K \|f\|_\infty ,$$

($K > 0$ is a constant independent of f).

Proof: Let $x_1, ..., x_L \in G$ be so that $|\tilde{f}(x_1, ..., x_L)| = \|\tilde{f}\|_\infty$

Let $B \subset \{1, ..., L\} \equiv \bar{L}$ and define for each $\gamma \in \Gamma$

$$\phi_B(\gamma) = \sum_{j \in B} \gamma(x_j) = \rho_B(\gamma) \exp(i\theta_B(\gamma)) .$$

$$(\rho_B(\gamma) = |\phi_B(\gamma)| \, and \, \theta_B(\gamma) = \arg\phi_B(\gamma))$$

By the symmetry of $\bar{F} \subset \Gamma^L$, observe that

$$\frac{1}{L!} |\tilde{f}(x_1, ..., x_L)| = \left| \sum_{m=1}^{L} (-1)^m \sum_{\substack{B \subset L \\ |B| = m}} \left(\sum_{\gamma \in F} a_\gamma \phi_B(\pi_1(\gamma)) ... \phi_B(\pi_L(\gamma)) \right) \right| . \quad (57)$$

Fix $B \subset \bar{L}$, and write the Riesz product

$$\mu_B = \prod_{\gamma \in \mathrm{E}} \left(1 + \frac{\rho_B(\gamma)}{|B|} (\exp(i\theta_B(\gamma))\gamma + \exp(-i\theta_B(\gamma))\bar{\gamma} / 2 \right) .$$

We have

$$\left| \sum_{\gamma \in \bar{F}} a_\gamma \phi_B(\pi_1(\gamma)) ... \phi_B(\pi_L(\gamma)) \right| = L! \left| \sum_{\gamma \in \bar{F}} \hat{f}(\gamma) \phi_B(\pi_1(\gamma)) ... \phi_B(\pi_L(\gamma)) \right| = L!(2|B|)^L |\mu_B * f(0)| \le L!(2|B|)^L \|f\|_\infty \quad (58)$$

Summing (58) over all $B \subset \bar{L}$, we obtain via (57) the desired estimate

Theorem (5.1.27)[226]: Let $\mathrm{E}\subset\Gamma$ be a dissociate set, $F\subset[\mathrm{E}_L]$ be arbitrary, and $\tilde{F}$ be given by (51). Then:

(i) $\theta_F=\dim\tilde{F}$; (ii) $r_F=\frac{(\dim\tilde{F}-1)}{2}$; (iii) $\sigma_F=\frac{2}{\left(1+\frac{1}{\dim\tilde{F}}\right)}$; (iv) $\eta_F=\frac{\left(1-\frac{1}{\dim\tilde{F}}\right)}{2}$.

Moreover θ_F,r_F,σ_F and η_F are exact if and only if dim $\tilde{F}$ is exact.

Proof :(iv). Suppose $\dim\tilde{F}=\alpha$ exact. Let $n>0$ be an arbitrary integer, and $f\in C_F(G)$ be so that

$$\left|\left\{\gamma\in F:\hat{f}(\gamma)\neq 0\right\}\right|=n..$$

By Holder's inequality and Theorem (5.1.26) (iii) we have

$$\sum\left|\hat{f}(\gamma)\right|\le\left\|\hat{f}\right\|_p n^{1/q}\le K\|f\|_\infty n^{1/q}$$

wheneverp $p\ge 2/(1+1/\alpha)$, and $1/p+1/q=1$. Therefore,

$$\Phi_F(n)\le Kn^a \text{ for all } a\ge(1-1/\alpha)/2 \tag{59}$$

We now recall the following basic fact:

An immediate corollary to Theorem (5.1.27) yields that the dimension of a spectral set $F\subset[\mathrm{E}]_L\subset\Gamma$is well defined. Suppose that E_1 and E_2 are dissociate sets and L_1, L_2 are positive integers. Assume that $F\subset[\mathrm{E}_1]_{L_1}$and $F\subset[\mathrm{E}_2]_{L_2}$. Denoting the images of F in $[\tilde{\mathrm{E}}_1]_{L_1}$ and $[\tilde{\mathrm{E}}_2]_{L_2}$ (given by(51)) as $\tilde{F}_1$ and$\tilde{F}_2$, respectively, we have

Corollary (5.1.28)[226]:

$$\dim\tilde{F}_1=\dim\tilde{F}_2=\dim F .$$

Moreover, dim $\tilde{F}_1$ is exact if and only if dim $\tilde{F}_2$ is exact. Theorem (5.1.27) is essentially a summary of the results in this section ,we shall sketch its proof in the case $F\subset\mathrm{E}_L$, and then indicate how to obtain the general case $F\subset[\mathrm{E}]_L$

.

Combining Lemmas (5.1.24) and (5.1.25) we deduce part (i) in Theorem . (5.1.27) Parts (ii) and (iii) in Theorem(5.1.27) follow from Theorems (5.1.11) and (5.1.16)thru symmetrizing procedures and the use of Riesz products similar to the ones employed in the proof of part (i). Leaving the details to the reader, we note that the added ingredient here is a simple combinatorial device, Lemma (5.1.26) below. Let $f \in C_F(G)$be a trigonometric polynomial, $f = \sum_{\gamma \in F} \hat{f}(\gamma)\gamma$, and define the trigonometric polynomial $\tilde{f} \in C_{\bar{F}}(G^L)$by $\hat{f} = \sum_{\gamma \in \tilde{F}} a_\gamma(\pi_1(\gamma), ..., \pi_L(\gamma))$, where $a_\gamma = \hat{f}(\pi_1(\gamma)...\pi_L(\gamma))$

for each $\gamma \in \bar{F}$ ($\bar{F}$ is defined by (53) at the outset of the proof of Lemma (5.1.26)**.**

Theorem (5.1.29) Let $F \subset \Gamma$ be a Sidon set with Sidon constant K_F. Then, for all $f \in L_F^2(G)$ and $2 < p < \infty$

$$K_F\sqrt{p}\|f\|_2 \geq \|f\|_p$$

Let $A \subset \text{E}$ be as at the start of the proof of Lemma (5.1.25) whence (as per the proof of Lemma (5.1.25)

$$\lambda_A L_{\cap F}(2s) \geq 2^{-L-1}(\Psi_F(s))^{1/2}$$

Therefore, by the Theorem above and the definition of Φ_F we have

$$K_A L_{\cap F} \geq 2^{2-L-1}\left(\frac{\Psi_F(s)}{2s}\right)^{1/2}.$$

From the inequality above it follows that

$$\overline{\lim_n} \frac{\Phi_F(n)}{n^a} < \infty \tag{60}$$

only if $a \geq \frac{\left(1 - \frac{1}{a}\right)}{2}$.

Combining (59) and (60), we obtain Theorem (5.1.27) (iv). The same proof works for the case $\dim \tilde{F} = a$asymptotic.
∎

The passage from $F \subset \mathrm{E}^L$ to $F \subset [\mathrm{E}]_L$ in Theorem (5.1.27) is based on the following basic lemma whose proof rests on routine Riesz product arguments.

Lemma (5.1.30) [226]: Let $1 \le k \le L$ and $\varepsilon = (\varepsilon_j)_{j=1}^k, \varepsilon_j = \pm 1,$ be arbitrary. There is $\mu \in M(G)$ so that

$$\hat{\mu}(\gamma) = \begin{cases} 1 & if\ \gamma \in \mathrm{E}_k^\varepsilon \\ 0 & if\ \gamma \in [\mathrm{E}]_L \setminus \mathrm{E}_k^\varepsilon. \end{cases}$$

Consequences ;Let E be a maximal dissociate set in Γ in which case we clearly have

$$\Gamma = \bigcup_{L=1}^{\infty} [\mathrm{E}]_L$$

and $[\mathrm{E}]_{L+1} \supset [\mathrm{E}]_L$,for all $L > 0$. The results in [231] can be translated to our present context to fill the 'gaps' between $[\mathrm{E}]_L$ and $[\mathrm{E}]_{L+1}$, $L = 1, \ldots.$.

Theorem (5.1.31)[226]: ([231]). Let $\mathrm{E} \subset \Gamma$ be a maximal dissociate set in Γ

(a) there exists a family of sets $\{F_x\}_{x \in [1,\infty)}$ with the following properties:

(i) For each $x \in 1, \infty), F_x \subset [\mathrm{E}]_{[x]}$ ($[x]$ denotes the smallest integer greater than x), and $\dim F_x = x$.

(ii) When $L < x < L+1$, L a positive integer, $F_x = \bigcup_{t<x} F_t$, and $F_L = [\mathrm{E}]_L = \bigcup_{t \le L} F_t$..In particular,

$$\bigcup_{x \in [1,\infty)} F_x = \Gamma.$$

(b) Let $x_0 \in [1, \infty)$ be arbitrary. There exists a family of sets $\{F_x\}_{x \in [1, x_0)}$ with the following properties:

(iii) For each $x \in [1, x_0), F_x \subset [\mathrm{E}]_{[x]}$ and $\dim F_x = x$ asymptotically. (iv) For each $x \in [1, x_0), F_x = \bigcap_{x_0 > t > x} F_t$.

Predictably, combining Theorem (5.1.31) (1) with Theorem (5.1.27) we obtain that Γ is a 'continuous' union of spectral sets whose combinatorial and analytic complexities are 'continuously' indexed. Similarly, part (2) of Theorem(5.1.31) yields the existence of continuously decreasing towers of asymptotic spectral sets whose combinatorial and analytic complexities are continuously indexed as well. Theorem**.** (5.1.16) implies, in effect, a statement that is stronger than part (ii) of Theorem (5.1.31)

Theorem (5.1.32)[226]: Let $E \subset \Gamma$ be dissociate, L a positive integer and $F \subset [E]_L$. Then: For all $f \in C_{|E|_L}(G)$

$$\left\| \hat{f} X_F \right\|_p < \infty$$

if and only if

$$\begin{cases} p \geq \dfrac{2}{\left(1+\dfrac{1}{\dim F}\right)} \dim F \; exact, \\ p \geq \dfrac{2}{\left(1+\dfrac{1}{\dim F}\right)} \dim F \; asymptotic. \end{cases}.$$

Section(5.2): Combinatorial Measurements and Orlicz Norms

We focus on connections between measurements reflecting purely combinatorial data and measurements that are based on harmonic-analytic and probabilistic properties.

Given an infinite set Y and $F \in Y^n (n \geq 1)$, we consider a function associated with $\Psi_F : \mathrm{N} \to \mathrm{N}$ such that for $s \in \mathrm{N}$,

$$\Psi_F(s) = \max\left\{ \left| F \cap (A_1 \times \ldots \times A_n) \right| : A_j \subset Y, \left| A_j \right| \leq s, \; j = 1, \ldots, n \right\}. \tag{61}$$

Define

$$\dim F = \overline{\lim_{s \to \infty}} \log \Psi_F(s) / \log s; \tag{62}$$

Equivalently , for $a > 0$ define

$$d_F(a) = \sup\left\{ \Psi_F(s) / s^a : s \in N \right\} \tag{63}$$

and observe that if $|F| = \infty$, then

$$\dim F = \inf\left\{ a : d_F(a) < \infty \right\} = \sup\left\{ a : d_F(a) = \infty \right\}. \tag{64}$$

The function Ψ_F is viewed as a gauge of the combinatorial complexity in

$F : \Psi_F(s)$ is the smallest integer K such that for all s-sets $A_1 \subset Y, \ldots, A_n \subset Y$ the number of samplings

$a_1 \in A_1, \ldots, a_n \in A_n$ with $(a_1, \ldots, a_n) \in F$ is no greater than k. The index dim is viewed as the combinatorial dimension of F, conveying that $\Psi_F(s)$ "grows like" $s^{\dim F}$, in the sense that

$$\overline{\lim_{s\to\infty}} \frac{\Psi_F(s)}{s^{\beta}} = \begin{cases} 0 & if\ \beta > \dim F, \\ \infty & if\ \beta < \dim F. \end{cases} \tag{65}$$

We distinguish between two cases:

(i) If $\overline{\lim}_{s\to\infty} \Psi_F(s)/s^{\dim F} < \infty (d_F\ (\dim F) < \infty)$, then $\dim F$ is exact;

(ii) If $\overline{\lim}_{s\to\infty} \Psi_F(s)/s^{\dim F} = \infty (d_F\ (\dim F) = \infty)$, then $\dim F$ is asymptotic.

(see [219]). In this section we further analyze the asymptotic case, and establish a precise resolution of it.

We take Y to be N (without loss of generality), and identify it with the Rademacher system $(r_j)_{j\in N} := \mathbf{R}$, a set of projections from $\{-1,1\}^N := \Omega\ onto \{-1,1\}$:

$$r_j(w) = w(j), \quad j \in N, \quad w = (w(j))_{j\in N} \in \Omega. \tag{66}$$

Here we view Ω as a compact Abelian group (endowed with the product topology, coordinate wise multiplication, and the normalized Haar measure P), and view $\mathbf{R}$ as an independent set of characters on Ω. (see [219]).For $F \subset \mathbf{R}^n (n \geq 1)$. , let $C_F(\Omega^n)$ and $L_F^2(\Omega^n)$ be, respectively, the spaces of continuous functions and P^n-square integrable functions on Ω^n, whose Fourier–Walsh transforms are supported in F.

For $t > 0$, let $\|\cdot\|$ be the ℓ^t norm, and for $f \in C(\Omega^n)$., let $\hat{f}$ be the Fourier-Walsh transform of f . For $F \subset \mathbf{R}^n$ and $t > 0$, let

$$\zeta_F(t) = \sup\left\{\left\|\hat{f}\right\|_t : f \in B_{C_F}(\Omega^n)\right\}, \tag{67}$$

where $B_{C_F}(\Omega^n)$ denotes the closed unit ball in $C_F(\Omega^n)$, and define

$$\sigma_F = \inf\{t : \zeta_F(t) < \infty\} = \sup\{t : \zeta_F(t) = \infty\}; \tag{68}$$

if $\zeta_F(\sigma_F) < 0$, then σ_F is exact, and if $\zeta_F(\sigma_F) = 0$, then σ_F is asymptotic.

For $F \subset \mathbf{R}^n$ and t> 0 , let

$$\eta_F(t) = \sup\left\{\|f\|_{LP} / p^t : p > 2, f \in B_{L_F^2(\Omega^n)}\right\}, \tag{69}$$

where $B_{L_F^2(\Omega^n)}$ is the closed unit ball in $L_F^2(\Omega^n)$ and define

$$\delta_F = \inf\{t : \eta_F(t) < 0\} = \sup\{\eta_F(t) = \infty\} \tag{70}$$

again, if $\eta_F(\delta_F) < 0$, then δ_F is exact, and if $\eta_F(\delta_F) < 0$, then δ_F is asymptotic.

The main results in [226] were:

$$d_F(t) < \infty \Leftrightarrow \zeta_F(2t/(t+1)) < \infty \Leftrightarrow \eta_F(t/2) < \infty \tag{71}$$

In particular,

$$\sigma_F = \frac{2\dim F}{\dim F + 1} \tag{72}$$

and

$$\delta_F = \frac{\dim F}{2}, \tag{73}$$

where σ_F and δ_F are exact if and only if $\dim F$ is exact. These results in effect were extensions of the classical Littlewood $2\mathrm{n}/(\mathrm{n}+1)$-inequalities [84,112], and the n-dimensional Khintchin inequalities [13,10].

We use Orlicz functions and their associated Orlicz norms to precisely resolve the case $d_F(\dim F) = \infty$. Our work is divided into four parts. In the first part we focus on the combinatorial gauge $\Psi_F, F \subset \mathbf{R}^n (n \geq 1)$. Given functions $\Psi : N \to N$ and $\Phi : R \to R$ we say that Ψ is quasi-asymptotic to Φ, and write $\Psi :\ _q \Phi$, if

$$0 < \overline{\lim_{s\to\infty}} \frac{\Psi(s)}{\Phi(s)} < \infty. \tag{74}$$

We prove Theorem (5.2.3) that if $F \subset \mathbf{R}^n$ is infinite, $\dim F = \alpha (\alpha \geq 1)$, and $\overline{\lim}_{s\to\infty} \Psi_F(s)/s^\alpha > 0$, then there exists an α-Orlicz function (Definition (5.2.1)) Φ such that $\Psi_F :\ q\Phi$.. Conversely, we show (Theorem (5.2.7) that for every α-Orlicz function $\Phi(\alpha \geq 1)$ there exists $F \subset \mathbf{R}^n$ such that $\Psi_F :\ q\Phi$. These results extend prior constructions in [227] and [220].

In the next three parts we derive precise relations between $\Psi_F\left(F \subset \mathbf{R}^n\right)$ and corresponding Orlicz norms associated with Ψ_F in $C_F\left(\Omega^n\right)$ and $L_F^2\left(\Omega^n, P^n\right)$ (Theorem(5.2.13) Corollary (5.2.18)and Theorem (5.2.22).These results naturally extend prior results stated in (71), (72) and (73) above, concerning relations between combinatorial dimension and Littlewood-type inequalities and Khintchin-type inequalities. An R -valued function Φ on $[0,\infty)$ is an Orlicz function if Φ is continuous, non-decreasing, convex, Φ (0) = 0, and $\lim_{x\to\infty}\Phi(x)=\infty$, see [116]. For $F \subset N^2$, and Orlicz function Φ, define (extending the definition in (63))

$$d_F(\Phi)=\sup\left\{\Psi_F(s)/\Phi(s): s\in N\right\}, \tag{75}$$

If $\Phi(x)=x^n$ for some $a \geq 1$, then we write $d_F(a)$ for $d_F(\Phi)$.

Note that $\dim F=\alpha$ is exact $(\alpha \geq 1)$ and $\overline{\lim}_{s\to\infty}\psi_F(s)/s^\alpha > 0$ if and only if Ψ_F is quas-iasymptotic to $\Phi(x)=x^\alpha, x\geq 0$. If $\dim F=\alpha$ is asymptotic, then we focus on $\phi(s)=\Psi_F(s)/s^\alpha$, where (necessarily) $\overline{\lim}_{s\to\infty}\phi(s)$, and $\phi(s)$ is $o\left(s^\varepsilon\right)$ for all $\varepsilon>0$. To this end, for technical reason that will later become apparent, we introduce the notion of an α-Orlicz function:

Definition (5.2.1.)[218]: For $\alpha \geq 1$, an Orlicz function Φ is said to be an α-Orlicz function if $\phi \in C^2[0,1)$ and $\Phi(x)=x^\alpha\phi(x)$ for $x\geq 0$, where either $\phi\equiv 1$, or ϕ satisfies the following properties:

(i) ϕ is concave and strictly increasing to ∞;(ii) $x\phi(x)$ is convex for $x \geq 0$; (iii) $\phi(x)=o\left(x^\varepsilon\right)$ for all $\varepsilon>0$, and for each $\varepsilon>0$ there exists $K>0$, such that $\phi(x)/x^\varepsilon$ is decreasing with increasing x for $x\geq K$.

Example(5.2.2)[218]: Suppose we want to construct an α-Orlicz function whose graph contains $\left(s, s^\alpha(\log s)^\beta\right)$ for s large, for some $\alpha \geq 1$ and $\beta>0$. Note that $(\log x)^\beta$ is not concave for $x<e^{\beta-1}$, $x(\log x)^\beta$ is not convex for $x<e^{1-\beta}$, and the y-intercept of the tangent line to the graph of $(\log x)^\beta$ at x for $x<e^\beta$ is less than 0. Let $x_0=\max\left\{e^{1-\beta}, e^\beta\right\}+1$, and let ℓ be the linear function whose graph is the tangent line to the graph of $(\log x)^\beta$ at x_0; that is

$$\ell(x)=\left(\log x_0\right)^\beta+\beta x_0^{-1}\left(\log x_0\right)^{\beta-1}\left(x-x_0\right), \qquad -\infty<x<\infty \tag{76}$$

Let

$$\tilde{\phi}(x)=\begin{cases}(\log x)^{\beta} & if\ x \geq x_0 \\ \ell(x) & if\, 0 \leq x < x_0\end{cases} \tag{77}$$

Smooth $\tilde{\phi}$ at x_0 so that the smoothed function ϕ is in $C^2[0,\infty)$, ϕ is concave, and $x\phi(x)$ is convex. Then the function $\Phi(x)=x^{\alpha}\phi(x)$ for $x \geq 0$ is the desired α-Orlicz function.

Theorem (5.2.3)[218]: Let $n \in \mathbb{N}$. If $F \subset \mathbb{N}^n$ is infinite with $\dim F=\alpha$, and $\overline{\lim}_{s\to\infty}\Psi_F(s)/s^{\alpha}>0$, then there exists an α-Orlicz function Φ such that $\Psi_F :\ q\Phi$..

Proof. Because $F \subset \mathbb{N}^2$ is infinite, we have $\alpha \geq 1$. If $d_F(\alpha)<\infty$, then $\Phi(x)=x^{\alpha}$ for $x \geq 0$ is an α-Orlicz function such that $\Psi_F :\ q\Phi$.

Suppose $d_F(\alpha)=\infty$. First we choose a sequence $\{s_j\}, s_j \uparrow \infty$. For any positive integers s and s', let $\ell_{s.s}$ be the linear function whose graph is the line passing through $(s, \Psi_F(s)/s^{\alpha})$ and $(s', \Psi_F(s')/(s')^{\alpha})$ let $\ell_{0,1}$ be the linear function whose graph is the line passing through (0, 0)and (1, 1).) Let $s_1=0$, and $s_2=1$. To choose for $j>2$, we proceed by (double) induction.

Suppose we have chosen s_j for $j>2$. To choose s_{j+1}, we consider the j points $s_j^{(1)},\dots,s_j^{(j)}$ such that

$$s_j^{(1)}=\min\left\{s>s_j : \frac{\Psi_F(s_j)}{s_j^{\alpha}}<\frac{\Psi_F(s)}{s^{\alpha}}<\ell_{s_{j-1},s_j}s\right\}, \text{and for } 1<i\leq j \tag{78}$$

$$s_j^{(1)}=\min\left\{s>s_j^{(i-1)} : \frac{\Psi_F(s_j)}{s_j^{\alpha}}<\frac{\Psi_F(s)}{s^{\alpha}}<\ell_{s_{j-1},s_j}s\right\}, \tag{79}$$

The existence of $s_j^{(1)},\dots,s_j^{(j)}$ for any j is guaranteed because $d_F(\alpha)=\infty$, and because $\Psi_F(s)/s^{\alpha}=o(s^{\varepsilon})$ for all $\varepsilon>0$ (because $\dim F=\alpha$). Denote the slope of _ $\ell_{s,s'}$ by $m_{s,s'}$ for any s and s'. Let

$$s_{j+1}=\max\left\{s\in\left[s_j^{(1)},\dots,s_j^{(j)}\right]: m_{s_j,s}\geq m_{s_j,s_j^{(1)}}\right\} \text{ for all } i=1,\dots,j. \tag{80}$$

Continuing this process, we obtain a sequence $s_j \uparrow \infty$ that satisfies

(i) $\Psi_F(s_j)/s_j^\alpha$ is strictly increasing to ∞ with increasing j ; (ii) $m_{s_{j-1},s_j} > m_{s_j,s_{j+1}} > 0$ for all $j > 1$;

(iii) for each j, and $s_j \le s \le s_j^{(j)}$ either

$$\frac{\Psi_F(s)}{s^\alpha} \ge \ell_{s_{j-1}s_j}(s), \tag{81}$$

or

$$\frac{\Psi_F(s)}{s^\alpha} \le \ell_{s_j,s_{j+1}}(s), \tag{82}$$

Claim (5.2.4)[218]: For each j , there are only finitely many $s \in \mathrm{N}$ such that

$$\frac{\Psi_F(s)}{s^\alpha} \ge \ell_{s_j,s_{j+1}}(s), \tag{83}$$

Proof : Suppose the claim is false. Then there exist j and a sequence $s'_k \uparrow \infty$ such that .

$$\frac{\Psi_F(s'_k)}{(s'_k)^\alpha} \ge \ell_{s_j,s_{j+1}}(s'_k), \tag{84}$$

For $x \ge 0$, write

$$\ell_{s_j,s_{j+1}}(x) = m_{s_j,s_{j+1}} x + b_j, \tag{85}$$

where $m_{s_j,s_{j+1}} > 0$, and $b_j > 0$. By (84) and (85),

$$\Psi_F(s'_k) \ge m_{s_j,s_{j+1}} (s'_k)^{a+1} + b_j (s'_k)^a \tag{86}$$

which contradicts $F = \alpha$, and the claim follows.
□

Let ℓ be the piecewise-linear function defined by

$$\ell(x) = \ell_{s_j s_{j+1}}(x), \qquad s_j \le x \le s_{j+1}, j \ge 1 \tag{87}$$

Claim (5.2.5)[218]:

$$\sup\left\{\Psi_F(s)/\left(s^\alpha \ell(s)\right) : s \in \mathrm{N}\right\} < \infty. \tag{88}$$

Proof : Suppose the claim is false. Then there exists a sequence $\{s_i'\}$ such that nd $\lim_{i\to\infty} \Psi_F\left((s_i')/(s_i')^\alpha \ell(s_i')\right) = \infty$.By Claim 1 and because $\left|s_j, s_j + j\right| \subset \left|s_j, s_j^{(j)}\right|$,], there exist j sufficiently large, and $s_i' \in \left[s_j, s_j^{(j)}\right]$ such that $\ell_{s_j, s_j^{(j)}}(s_i') < \Psi_F(s_i')/(s_i')^\alpha < \ell_{s_{j-1}, s_j}(s_i')$, which contradicts (81) and (82), and the claim follows.

Next we construct a spline function as follows. Note that for $b > 0, (\log x)^b$ is concave for $b < \log + 1$, and $x(\log x)^b$ is convex for $x > e$. We start from s_4 (because $s_4 > e$). For $s_4 \le x \le s_5$, let

$$p_4(x) = a_4(\log x)^{b_4} + c_4 x + d_4, \tag{89}$$

Where $a_4 > 0, 0 < b_4 < \log s_4 + 1, c_4 \ge 0$, and d_4 are chosen such that C

$$p_4(s_4) = \ell(s_4), \quad p_4(s_5) = \ell(s_5), \quad (p_4)'_+(s_4) = \frac{m_{s_3,s_4} + m_{s_4,s_5}}{2} \tag{90}$$

where $(p_4)'_+(x)$ denotes the right derivative of p_4 at x. (Similarly $(p_4)'_-(x)$ denotes the left derivative of p_4 at x.)

For $s_5 \le x \le s_6$, let

$$p_5(x) = a_5(\log x)^{b_5} + c_5 x + d_5. \tag{91}$$

where $a_5 > 0, 0 < b_5 < \log s_5 + 1, c_5 \ge 0$, and d_5 are chosen such that:

(iv) if $(p_4)'_-(s_5) > m_{s_5,s_6}$, then

$$p_5(s_5) = \ell(s_5), \quad p_5(s_6) = \ell(s_6), \quad \text{and}\, (p_5)'_+(s_5) = (p_4)'_-(s_5); \tag{92}$$

(v) if $(p_4)'_-(s_5) \le m_{s_5,s_6}$, then

$$p_5(s_5) = \ell(s_5), \quad p_5(s_6),\ and\ (p_5)'_-(s_6) = \frac{m_{s_5,s_6} + m_{s_6,s_7}}{2}. \tag{93}$$

We proceed as follows. For $j \ge 6$, and $s_j \le x \le s_{j+1}$, let

$$p_j(x) = a_j(\log x)^{b_j} + c_j x + d_j, \tag{94}$$

where $a_j > 0, 0 < b_j < \log s_j + 1, c_j \ge 0,$, and d_j are chosen such that:

(vi) if $\left(p_{j-1}\right)'_{-}\left(s_j\right) > m_{s_j,s_{j+1}}$, then

$$p_j\left(s_j\right)=\ell\left(s_j\right),\quad p_j\left(s_{j+1}\right)=\ell\left(s_{j+1}\right),\text{ and }\left(p_j\right)'_{+}\left(s_j\right)=\left(p_{j-1}\right)'_{-}\left(s_j\right); \tag{95}$$

(vii) if $\left(p_{j-1}\right)'_{-}\left(s_j\right) \le m_{s_j,s_{j+1}}$, then

$$p_j\left(s_j\right)=\ell\left(s_j\right),\qquad p_j\left(s_{j+1}\right)=\ell\left(s_{j+1}\right), \tag{96}$$

and

$$\left(p_j\right)'_{-}\left(s_{j+1}\right)=\frac{m_{s_j,s_{j+1}}+m_{s_{j+1},s_{j+2}}}{2}. \tag{97}$$

For any $j \ge 5$ such that (vii) holds, $\left(p_{j-1}\right)'_{-}\left(s_j\right) \le m_{s_j,s_{j+1}} < \left(p_j\right)'_{+}\left(s_j\right)$. By the mean value theorem, there exist $x_{j-1} \in \left(s_{j-1},s_j\right)$, and $x_j \in \left(s_j,s_{j+1}\right)$ such that $p_{j-1} \in \left(x_{j-1}\right) = m_{s_{j-1},s_j}$ and $p'_j \in \left(x_j\right) = m_{s_j,s_{j+1}}$ Because p_{j-1} and p_j are concave, and because $m_{s_{j-1},s_j} > m_{s_j,s_{j+1}}$, there are $t_{j-1} \in \left(x_{j-1},s_j\right)$, , and $t_j \in \left(s_j,x_j\right)$ such that

$$p'_{j-1}\left(t_{j-1}\right)=p'_j\left(t_j\right). \tag{98}$$

For $x \ge 0$, let

$$T_j\left(x\right)=p_{j-1}\left(t_{j-1}\right)+p'_{j-1}\left(t_{j-1}\right)\left(x-t_{j-1}\right), \tag{99}$$

that is, T_j is the linear function whose graph is both the tangent line to the graph of p_{j-1} at t_{j-1},and the tangent line to the graph of p_j at t_j .Let $\tilde{\phi}$ be the spline function such that

(viii) for $0 \le x \le s_4, \tilde{\phi}\left(x\right)$ is the linear function whose graph is the tangent line to the graph of p_4 at s_4;

(x) for any $x \ge s_4$, let $p\left(x\right)=p_j\left(x\right), s_{j-1} \le x \le s_j$ for $j \ge 5$ and let

$$\tilde{\phi}\left(x\right)=\begin{cases} T_j\left(x\right) & \text{if } t_{j-1} \le x \le t_j, [t_{j-1},t_j] \subset [x_{j-1},x_j],\ j \ge 5, \\ p\left(x\right) & \text{otherwise,} \end{cases} \tag{100}$$

where t_{j-1} and $t_{j,}j \ge 5$, are indicated in (98).

Then $\tilde{\phi}$ is concave, and $x\tilde{\phi}\left(x\right)$ is convex. Let $\Phi = x^{\alpha}\tilde{\phi}\left(x\right)$. By Claim (5.2.5) and because $\tilde{\phi} \ge \ell$,

$$\sup\left\{\Psi_F(s)/\Phi(s): s\in\mathrm{N}\right\}\le\sup\left\{\Psi_F(s)/\left(s^{\alpha}\ell(s)\right): s\in\mathrm{N}\right\}<\infty . \tag{101}$$

Claim (5.2.6)[218]: There are infinitely many j such that

$$\tilde{\phi}(s_j)=\ell(s_j). \tag{102}$$

Next we establish the converse to Theorem (5.2.3.) .

Lemma (5.2.7)[218]:[219] Let $n\ge 2$ be an integer, and $1\le\gamma<n$. Let Φ be an Orlicz function such that $x\le\Phi(x)\le x^{\gamma}$ for all $x\in[1,\infty)$ and $\Phi(x)/x^{\gamma}$ is decreasing with increasing x. Then for every $k\in\mathrm{N}$, there exist $F\subset[k]^n\left([k]=\{1,...,k\}\right)$ such that

$$\Psi_F(s)\le C\Phi(s),\quad s\in[k], \tag{103}$$

and

$$|F|=\Psi_F(k)\ge\frac{1}{2}\Phi(k). \tag{104}$$

where C > 0 depends only on n and γ .

Proof: For k $k\in\mathrm{N}$, let $\left\{X_i^{(k)}: i\in[k]^n\right\}$ be the Bernoulli system of statistically independent {0, 1}-valued variable on (Ω,P) such that

$$\mathrm{P}\left(X_i^{(k)}=1\right)=\frac{\Phi(k)}{k^n}. \tag{105}$$

Consider the random set $F=\left\{i: X_i^{(k)}=1\right\}$.

We use the following elementary fact about binomial probabilities: for $p\in(0,1)$, and integers $m>0$ and $i\ge 2mp$,

$$2\binom{m}{i+1}p^{i+1}(1-p)^{m-i-1}\le\binom{m}{i}p^{i}(1-p)^{m-i}, \tag{106}$$

which implies

$$\sum_{i=j}^{m}\binom{m}{i}p^{i}(1-p)^{m-i}\le 2\binom{m}{j}p^{j}(1-p)^{m-j},\quad j\ge 2mp. \tag{107}$$

Fix $s\in[k]$, and let A be a s-hypercube in ($[k]^n$ $A=A_1\times\ldots\times A_n$ where $|A_1|=\ldots=A_n=s$) Denote

$$C=\max\left\{2e^{n+2},\frac{n+1}{n-\gamma}\right\}. \tag{108}$$

Let $j(s)=\lceil C\Phi(s)\rceil$ (= smallest integer $\geq C\Phi(s)$). Then

$$j\geq 2\Phi(s)=2s^n\frac{\Phi(k)}{k^n}\frac{\Phi(s)}{s^n}\frac{k^n}{\Phi(k)}\geq 2s^n\frac{\Phi(k)}{k^n} \tag{109}$$

(because $\Phi(s)/s^\gamma$ is decreasing) . By (107) and (109),

$$\mathrm{P}\left(\sum_{i\in A}X_i^{(k)}\geq j\right)=\sum_{i=j}^{s^n}\binom{s^n}{i}\left(\frac{\Phi(k)}{k^n}\right)^i\left(1-\frac{\Phi(k)}{k^n}\right)^{s^n-i}\leq 2\binom{s^n}{j}\left(\frac{\Phi(k)}{k^n}\right)^j\left(1-\frac{\Phi(k)}{k^n}\right)^{s^n-j}$$

$$\leq 2\binom{s^{nj}}{j\,!}\left(\frac{\Phi(k)}{k^n}\right)^j\leq\frac{2s^{nj}\Phi(k)^j}{\left(C\Phi(s)\right)^j e^{-j}k^{nj}} \tag{110}$$

Then,

$$\mathrm{P}\left(\sum_{i\in A}X_i^{(k)}\geq C\Phi(s)\,\text{for some s}-\text{hypercube A}\right)\leq\binom{k}{s}^n\frac{2s^{nj}\left(\Phi(k)\right)^j}{\left(C\Phi(s)\right)^j e^{-j}k^{nj}}\leq\frac{k^{ns}}{s^{ns}e^{-ns}}\frac{2s^{nj}\left(\Phi(k)\right)^j}{\left(C^j\Phi(s)\right)^j e^{-j}k^{nj}}$$

$$\leq 2\frac{e^{j+ns}}{\left(2e^{n+2}\right)^j}\left(\frac{s}{k}\right)^{nj-ns}\left(\frac{\Phi(k)}{\Phi(s)}\right)\leq 2\frac{1}{e^{n(j-s)+j}}\left(\frac{s}{k}\right)^{nj-ns}\left(\frac{k}{s}\right)^{\gamma j}\leq e^{-s}\left(\frac{s}{k}\right)^{(n-\gamma)j-ns}\leq e^{-s}\left(\frac{s}{k}\right)^{sj}\left(by\,(49)\right) \tag{111}$$

(because $\Phi(s)/s^\gamma, n(j-s)+j\geq j\geq s$ and $(n-\gamma)j\geq(n-\gamma)C\Phi(s)\geq(n+1)s$)

Hence,

$$\mathrm{P}\left(\sum_{i\in A}X_i^{(k)}\geq C\Phi(s)\;for\;some\;s-hypercube\;A,\,s\in[k]\right)\leq\sum_{s=1}^{k}e^{-s}\left(\frac{s}{k}\right)^s \tag{112}$$

Therefore,

$$\lim_{k\to\infty}\mathrm{P}\left(\sum_{i\in A}X_i^{(k)}\geq C\Phi(s)\right)=0\,. \tag{113}$$

By Chebyshev's inequality,

$$\mathrm{P}\left(\left|\sum_{i\in[k]^n} X_i^{(k)} - \Phi(k)\right| > \frac{\Phi(k)}{2}\right) \le \frac{Var\left(\sum_{i\in[k]^n} X_i^{(k)} - \Phi(k)\right)}{\left(\frac{\Phi(k)}{2}\right)^2} = \frac{4k^n Var\left(X_i^{(k)}\right)}{\Phi(k)^2} \le \frac{4k^n\left(\frac{\Phi(k)}{k^n}\right)\left(1-\frac{\Phi(k)}{k^n}\right)}{\Phi(k)^2}$$

$$\le \frac{4}{\Phi(k)}. \quad (114)$$

Hence

$$\lim_{k\to\infty} \mathrm{P}\left(\sum_{i\in[k]^n} X_i^{(k)} \le \frac{\Phi(k)}{2}\right) = 0 \quad (115)$$

By (113) and (115),

$$\lim_{k\to\infty} \mathrm{P}(F) \quad (116)$$

satisfies (103) and (104)=1

□

Let $\pi_1,\dots,\pi_n$ be the canonical projections from $\mathbb{N}^n$ onto $\mathbb{N}$. We say $F\subset\mathbb{N}^n$ and $G\subset\mathbb{N}^n$ are n-disjoint if $\pi_\ell(F)\cap\pi_\ell(G)=\phi$ for all $\ell=1,\dots,n$.

Lemma (5.2.8)[218]: ($Cf.$[219].). Suppose $F_j, j\in\mathbb{N},$ is a sequence of pairwise n-disjoint subsets of $\mathbb{N}^n$, and let $F=\bigcup_j F_j$. For an Orlicz function Φ, and for every $m\in\mathbb{N}$,

$$\sup\left\{\Psi_F(s)/\Phi(s): s\in[m]\right\} \le n\sup\left\{\Psi_{F_j}(s)/\Phi(s): s\in[m], j\in\mathbb{N}\right\}. \quad (117)$$

Proof: Let $m\in\mathbb{N}$ and let $s\in[m]$. For $A_1\times\dots\times A_n\subset\mathbb{N}^n$ such that $|A_1|\le s,\dots,|A_n|\le s,$ let

$$s_{i,j} = \left|\pi_i(F_j)\cap A_i\right|, \quad i\in[n], j\in\mathbb{N} \quad (118)$$

and

$$s_j = \max\left\{s_{i,j} : i\in[n]\right\}, \quad j\in\mathbb{N} \quad (119)$$

Then $$\sum_{j=1}^{\infty} s_{i,j} \le |A_i| \le s, \quad i \in [n] \tag{120}$$

Let

$$L = \sup\{\Psi_F(s)/\Phi(s) : s \in [m],\ j \in \mathrm{N}\}. \tag{121}$$

By (118), (119) and (121), for any $j \in \mathrm{N}$,

$$\left|F_j \cap (A_1 \times ... \times A_n)\right| = \left|F_j \cap \left(\left(\pi_1(F_j) \cap A_1\right) \times ... \times \left(\pi_n(F_j) \cap A_n\right)\right)\right| \le L\Phi(s_j) \tag{122}$$

then

$$\frac{\left|F \cap (A_1 \times ... \times A_n)\right|}{\Phi(s)} = \frac{\sum_{j=1}^{\infty}\left|F_j \cap (A_1 \times ... \times A_n)\right|}{\Phi(s)} \le \frac{L\sum_{j=1}^{\infty}\Phi(s_j)}{\Phi(s)}. \tag{123}$$

Because Φ is increasing,

$$\sum_{j=1}^{\infty}\Phi(s_j) \le \sum_{i=1}^{n}\sum_{j=1}^{\infty}\Phi(s_{i,j}) \le \sum_{i=1}^{n}\Phi\left(\sum_{j=1}^{\infty}s_{i,j}\right) \le n\,\Phi(s) \text{ by (119)} \tag{124}$$

(because Φ is convex) by (120).

By (123) and (114),

$$\frac{\left|F \cap (A_1 \times ... \times A_n)\right|}{\Phi(s)} \le nL \tag{125}$$

Theorem (5.2.9)[218]: ([219].For $n \ge 2$, and $1 \le \alpha < n$,, if Φ is an α-Orlicz function, then there exist $F \subset \mathrm{N}^n$ such that $\Psi_F \sim \mathrm{q}\Phi$,

Proof]: Let $\alpha < \gamma < n$, and let Φ be an α-Orlicz function. Then $x \le \Phi(x) \le x^{\gamma}$ for large x, and $\Phi(x)/x^{\gamma}$ is eventually decreasing. By Lemma (5.2.3) we produce a collection $\{F_j\}$ of pairwise n-disjoint subsets of N^n such that $\{\Psi_{F_j}(s)/\Phi(s) : s \in \mathrm{N}\} < C$, and for each $j \in \mathrm{N}$, we have $\left|\pi_\ell(F_j)\right| = j$ for $\ell \in [n], \left|F_j\right| \ge \Phi(j)/2$. Let $F = \cup_j F_j$, and apply Lemma (5.2.9)

Suppose $\Phi(x) = x^{\alpha}\phi(x)$ for $x \ge 0$ is an α-Orlicz function. Because φ is concave, increasing, and $\phi(0) \ge 0$, we have $\phi'(x) \le \phi(x)/x$ for all $x \ge 0$. Hence

$$0 \le \frac{\phi'(x)}{\phi(x)} x \le 1, \qquad x \ge 0. \tag{126}$$

Let

$$\Theta(x) = x^{\frac{\alpha+1}{2}} \left(\phi(x)\right)^{\frac{1}{2}}, \quad x \ge 0 \tag{127}$$

and

$$\theta(x) = \frac{1}{\phi\left(\Theta^{-1}(1/x)\right)} \qquad , x \ge 0 \tag{128}$$

Note that

$$\phi(x)\theta\left(1/\Theta(x)\right) = 1, \qquad x \ge 0 \tag{129}$$

For $x \ge 0$, define

$$M_\Phi(x) = x^{\frac{2\alpha}{\alpha+1}} \left(\theta(x)\right)^{\frac{1}{\alpha+1}} \tag{130}$$

Then

$$M'_\Phi(x) = \frac{1}{\alpha+1} x^{\frac{2\alpha}{\alpha+1}-1} \left(\theta(x)\right)^{\frac{1}{\alpha+1}} \left\{2\alpha + \frac{\theta'(x)}{\theta(x)} x\right\}. \tag{131}$$

and

$$M''_\Phi(x) = \frac{1}{\alpha+1} x^{\frac{2\alpha}{\alpha+1}-2} \left(\theta(x)\right)^{\frac{1}{\alpha+1}} \left\{\frac{2\alpha(\alpha-1)}{\theta(x)} + Dx\right\}. \tag{132}$$

where

$$D(x) = \frac{4\alpha}{\alpha+1} \frac{\theta'(x)}{\theta(x)} x - \frac{\alpha}{\alpha+1} \left(\frac{\theta'(x)}{\theta(x)} x\right)^2 + \frac{\theta''(x)}{\theta(x)} x^2. \tag{133}$$

We now establish that M_Φ is an Orlicz function. We will use the Orlicz norm associated with M_Φ.

Lemma (5.2.10)[218]: M_Φ (defined in (130)) is an Orlicz function. Moreover, except for the case $\Phi(x) = x$ for $x \ge 0$, we have $M'_\Phi(x) > 0$ and $M''_\Phi(x) > 0$ for $x > 0$.

Proof: It is obvious that $M'_\Phi(x)>0$ for $x>0$. Now we consider M''_Φ. Taking derivatives on both sides of (129), we have

$$\phi'(x)\theta\left(1/\Theta(x)\right)+\phi(x)\theta'\left(1/\Theta(x)\right)\left(1/\Theta(x)\right)'=0. \tag{134}$$

Hence

$$\frac{\phi'(x)}{\phi(x)}x+\frac{\theta'\left(1/\Theta(x)\right)}{\theta\left(1/\Theta(x)\right)}\left(1/\Theta(x)\right)'x=0. \tag{135}$$

By (127),

$$\left(1/\Theta(x)\right)'x=\frac{\alpha+1}{2}\left(1+E(x)\right)\left(1/\Theta(x)\right), \tag{136}$$

where

$$\mathrm{E}(x)=\left(\frac{1}{\alpha+1}\right)\frac{\phi'(x)}{\phi(x)}x. \tag{137}$$

Note $\alpha\geq 1$. By (135) and (136), and by substituting $1/\Theta(x)=y$, we have

$$\frac{\theta'(y)}{\theta(y)}y=\frac{2}{(\alpha+1)\left(1+E(x)\right)}\frac{\phi'(x)}{\phi(x)}x\leq\frac{\phi'(x)}{\phi(x)}x\leq 1 \text{ by (126)}. \tag{138}$$

Taking derivatives on both sides of (134), we have

$$\phi''(x)\theta\left(1/\Theta(x)\right)+2\phi'(x)\theta'\left(1/\Theta(x)\right)\left(1/\Theta(x)\right)'$$
$$+\phi(x)\theta''\left(1/\Theta(x)\right)\left(1/\Theta(x)'\right)^2+\phi(x)\theta'\left(1/\Theta(x)\right)\left(1/\Theta(x)\right)''=0. \tag{139}$$

Hence

$$\frac{\phi''(x)}{\phi(x)}x^2+2\frac{\phi'(x)}{\phi(x)}x\frac{\theta'\left(1/\Theta(x)\right)}{\theta\left(1/\Theta(x)\right)}\left(1/\Theta(x)\right)'x\frac{\theta''\left(1/\Theta(x)\right)}{\theta\left(1/\Theta(x)\right)}\left(\left(1/\Theta(x)\right)'\right)^2x^2$$
$$+\frac{\theta'\left(1/\Theta(x)\right)}{\theta\left(1/\Theta(x)\right)}\left(1/\Theta(x)\right)''x^2=0 \tag{140}$$

By (127),

$$\left(1/\Theta(x)\right)''x^2=\frac{(\alpha+1)(\alpha+3)}{4}\left(1+F(x)\right)\left(1/\Theta(x)\right), \tag{141}$$

where

$$F(x)=\frac{4}{(\alpha+1)(\alpha+3)}\left\{\frac{\alpha+1}{2}\frac{\phi'(x)}{\phi(x)}x+\frac{3}{4}\left(\frac{\phi'(x)}{\phi(x)}x\right)^2-\frac{1}{2}\frac{\phi''(x)}{\phi(x)}x^2\right\}. \quad (142)$$

Bringing (136) and (141) into (140), and substituting $1/\Theta(x)=y$, we have

$$\frac{\phi''(x)}{\phi(x)}x^2-(\alpha+1)(1+\mathrm{E}(x))\frac{\phi'(x)}{\phi(x)}x\frac{\theta'(y)}{\theta(y)}y$$

$$+\left(\frac{\alpha+1}{2}\right)^2(1+\mathrm{E}(x))^2\frac{\theta''(y)}{\theta(y)}y^2\frac{(\alpha+1)(\alpha+3)}{4}(1+\mathrm{E}(x))\frac{\theta'(y)}{\theta(y)}y=0. \quad (143)$$

By (137) and (142),

$$1+F(x)=(1+\mathrm{E}(x))^2-\frac{2}{(\alpha+1)(\alpha+3)}\frac{\phi''(x)}{\phi(x)}x^2-G(x), \quad (144)$$

where

$$G(x)=\frac{4}{(\alpha+1)(\alpha+3)}\frac{\phi'(x)}{\phi(x)}x\left\{-\frac{\alpha}{2(\alpha+1)}\frac{\phi'(x)}{\phi(x)}x\right\}. \quad (145)$$

Then by (126), $G(x)\geq 0$ for all $x\geq 0$. Applying (138) and (144) to (143), we have

$$\frac{\phi''(x)}{\phi(x)}x^2-\frac{(\alpha+1)^2(1+\mathrm{E}(x))^2}{2}\left(\frac{\theta'(y)}{\theta(y)}y\right)^2$$

$$+\frac{(\alpha+1)^2(1+\mathrm{E}(x))^2}{4}\frac{\theta''(y)}{\theta(y)}y^2+\frac{(\alpha+1)(\alpha+3)}{4}(1+\mathrm{E}(x))^2\frac{\theta'(y)}{\theta(y)}y$$

$$-\frac{\phi''(x)}{2\phi(x)}x^2\frac{\theta'(y)}{\theta(y)}y-\frac{(\alpha+1)(\alpha+3)}{4}G(x)\frac{\theta'(y)}{\theta(y)}y=0. \quad (146)$$

Then

$$\frac{(\alpha+1)^2(1+\mathrm{E}(x))^2}{4}\left\{-2\left(\frac{\theta'(y)}{\theta(y)}y\right)^2+\frac{\theta''(y)}{\theta(y)}y^2+\frac{\alpha+3\theta'(y)}{\alpha+1\theta(y)}y\right\}$$

$$=\frac{\phi''(x)}{\phi(x)}x^2\left(1-\frac{1}{2}\frac{\theta'(y)}{\theta(y)}y\right)+\frac{(\alpha+1)(\alpha+3)}{4}G(x)\frac{\theta'(y)}{\theta(y)}y\geq 0$$

$$\left(because-\frac{\phi''(x)}{\phi(x)}x^2\geq 0,0\leq\frac{\theta'(y)}{\theta(y)}y\leq 1\, and\; G(x)\geq 0\right). \tag{147}$$

Hence for all $y\geq 0$.

$$-2\left(\frac{\theta'(y)}{\theta(y)}y\right)^2+\frac{\theta''(y)}{\theta(y)}y^2+\frac{(\alpha+3)\theta'(y)}{(\alpha+1)\theta(y)}y\geq 0. \tag{148}$$

Then, by (133), for all $x\geq 0$,,

$$D(x)=\left\{-2\left(\frac{\theta'(x)}{\theta(x)}x\right)^2+\frac{\theta''(x)}{\theta(x)}x^2+\frac{(\alpha+3)\theta'(x)}{(\alpha+1)\theta(x)}x\right\}.$$

$$+\frac{\alpha+2}{\alpha+1}\left(\frac{\theta'(x)}{\theta(x)}x\right)^2+\frac{3\alpha-3}{\alpha+1}\frac{\theta'(x)}{\theta(x)}x\geq 0 \tag{149}$$

By (132) and (149), we have $M_\Phi^n\geq 0$. If $\alpha>1$, then $2\alpha(\alpha-1)(\alpha+1)>0$, and hence $M_\Phi^n(x)>0$ for $x>0$. If $\alpha=1$, and φ is strictly increasing, then $\frac{\theta'(y)}{\theta(y)}y>0$ in (138), and hence $D(x)>0$ for $x>0$. Then $M_\Phi^n(x)>0$ for $x>0$. Because Φ is an α-Orlicz function, either $\phi\equiv 1$, or ϕ is strictly increasing. Therefore, except for the case $\Phi(x)=x$, we have $M'_\Phi(x)>0$ and $M''_\Phi(x)>0$ for $x>0$.

The following property will be needed.

Lemma (5.2.11)[218]: For M_Φ defined in (130), $M'_\Phi(x)-xM''_\Phi(x)>0$ for all $x>0$.

Proof: It suffices to show that $M'_\Phi(y)-yM''_\Phi(y)>0$ for all $y>0$, where $y=1/\Theta(x)$, For simplicity, we denote M_Φ by M_Φ. By (131) and (132),

$$M'(y)-yM''(y)=\frac{1}{\alpha+1}y^{\frac{2\alpha}{\alpha+1}-1}\theta(y)^{\frac{1}{\alpha+1}}H(y), \tag{150}$$

where

$$H(y) = 2\alpha + \frac{\theta'(y)}{\theta(y)} y - \frac{2\alpha(\alpha-1)}{\alpha+1} - Dy, \tag{151}$$

where $D(\mathrm{y})$ is defined in (133). By (149) and (147), we have

$$D(y) = \frac{4}{(\alpha+1)^2 (1+\mathrm{E}(x))^2} \left\{ -\frac{\phi''(x)}{\phi(x)} x^2 \left(1 - \frac{1}{2}\frac{\theta'(y)}{\theta(y)}\right) + \frac{(\alpha+1)(\alpha+3)}{4} G(x) \frac{\theta'(y)}{\theta(y)} y \right\}$$

$$+ \frac{(\alpha+2)}{(\alpha+1)} \left(\frac{\theta'(y)}{\theta(y)} y\right)^2 + \frac{3\alpha - 3\,\theta'(y)}{(\alpha+1)\theta(y)} y \tag{152}$$

Because $x\phi(x)$ is convex for $x \geq 0$, we have

$$(x\phi(x))'' = 2\phi'(x) + x\phi''(x) \geq 0, \quad x \geq 0. \tag{153}$$

Hence

$$-\frac{\phi''(x)}{\phi(x)} x^2 \leq 2\frac{\phi'(x)}{\phi(x)} x$$

(154)

By (151), (152), (154) and (145),

$$H(y) \geq \frac{4\alpha}{\alpha+1} + \frac{\theta'(y)}{\theta(y)} y$$

$$-\frac{4}{(\alpha+1)^2 (1+E(x))^2} \left\{ 2\frac{\phi'(x)}{\phi(x)} x \left(1 - \frac{1}{2}\frac{\theta'(y)}{\theta(y)} y\right) + \frac{\phi'(x)}{\phi(x)} x \left(1 - \frac{\alpha}{2(\alpha+1)} \frac{\phi'(x)}{\phi(x)} x\right) \frac{\theta'(y)}{\theta(y)} y \right\}$$

$$-\frac{\alpha+2}{\alpha+1} \left(\frac{\theta'(y)}{\theta(y)}\right)^2 - \frac{3\alpha-3}{\alpha+1} \frac{\theta'(y)}{\theta(y)} y$$

$$= \frac{4\alpha}{\alpha+1} + \frac{\theta'(y)}{\theta(y)} y - \frac{4}{(\alpha+1)(1+E(x))} \frac{\theta'(y)}{\theta(y)} y + \frac{\alpha}{2(\alpha+1)} \left(\frac{\theta'(y)}{\theta(y)} y\right)^3$$

$$-\frac{\alpha+2}{\alpha+1} \left(\frac{\theta'(y)}{\theta(y)} y\right)^2 - \frac{3\alpha-3}{\alpha+1} \frac{\theta'(y)}{\theta(y)} \quad \text{by (138)} \tag{155}$$

By (137),

$$\frac{4}{(\alpha+1)(1+E(x))}\frac{\theta(y)}{\theta(y)}y=\frac{4}{\alpha+1+\frac{\phi'(x)}{\phi(x)}x}\frac{\theta'(y)}{\theta(y)}y$$

$$\leq\frac{4}{\alpha+1+\frac{\phi'(x)}{\phi(x)}x}\frac{\phi'(x)}{\phi(x)}x\leq\frac{4}{\alpha+1}\quad \text{(by (126) and (78)) .} \tag{156}$$

By (155) and (156),

$$H(y)\geq\frac{4\alpha}{\alpha+1}-\frac{4}{\alpha+2}+\left(1-\frac{3\alpha-3}{\alpha+1}\right)\frac{\theta'(y)}{\theta(y)}y-\frac{\alpha+2}{\alpha+1}\left(\frac{\theta'(y)}{\theta(y)}y\right)^2$$

$$=\frac{\alpha^2}{(\alpha+1)(\alpha+2)}+\frac{3\alpha-2}{\alpha+1}+\frac{-2\alpha+4}{\alpha+1}\frac{\theta'(y)}{\theta(y)}y-\frac{\alpha+2}{\alpha+1}\left(\frac{\theta'(y)}{\theta(y)}y\right)^2$$

$$\geq\frac{\alpha^2}{(\alpha+1)(\alpha+2)}+\left(\frac{3\alpha-2}{\alpha+1}+\frac{-2\alpha+4}{\alpha+1}\right)\frac{\theta'(y)}{\theta(y)}y-\frac{\alpha+2}{\alpha+1}\left(\frac{\theta'(y)}{\theta(y)}y\right)^2$$

(because $0\leq\frac{\theta'(y)}{\theta(y)}y\leq1$)

$$=\frac{\alpha^2}{(\alpha+1)(\alpha+2)}+\frac{\alpha+2}{\alpha+1}\frac{\theta'(y)}{\theta(y)}y-\frac{\alpha+2}{\alpha+1}\left(\frac{\theta'(y)}{\theta(y)}y\right)^2>0, \tag{157}$$

as desired.

We recall the following definitions of Orlicz norms (see[116]). For an Orlicz function M and a sequence of scalars $a=(a_1,a_2,...)$, define

$$\|a\|_M=\inf\left\{\rho>0:\sum_{n=1}^{\infty}M(|a_n|/\rho)\leq1\right\}, \tag{158}$$

$$M^*(\mu)=\max\{\mu x-M(x):x>0\}, \tag{159}$$

and

$$|||a|||_M=\sup\left\{\sum_{n=1}^{\infty}a_nb_n:\sum_{n=1}^{\infty}M^*(|b_n|)\leq1\right\}. \tag{160}$$

The two Orlicz norms $\|\cdot\|_M$ and $|||\cdot|||_M$ are equivalent and

$$\|a\|_M \le \||a\||_M \le 2\|a\|_M . \tag{161}$$

Definition (5.2.12)[218]: For $F \subset \mathbf{R}^n$ and α-Orlicz function Φ, let

$$\zeta_F(\Phi) = \sup\left\{\left\|\hat{f}\right\|_{M_\Phi} : f \in B_{C_F}(\Omega^n)\right\}, \tag{162}$$

where M_Φ is given in (130).

This definition naturally extends the definition in (67). If $\Phi(x) = x^\alpha, x \ge 0$, for some $\alpha \ge 1$, then $\zeta_F(\Phi)$ and $\zeta_F(2\alpha/(\alpha+1))$ have the same meaning.

Let $n \in \mathrm{N}$. For $F \subset \mathbf{R}^n$ and α-Orlicz function Φ, let

$$\delta(\alpha) = \begin{cases} \dfrac{1}{2\alpha} & \text{if } d_F(\Phi) \le 1 \\ \dfrac{\alpha}{2(\alpha^2+\alpha+1)} & \text{if } d_F(\Phi) > 1 \end{cases} \tag{163}$$

Theorem (5.2.13)[218]:. (Cf.[219]) For $n \in \mathrm{N}$, there exist $C_n > 0$ and $D_n > 0$ such that for all $F \subset \mathbf{R}^n$ and α-Orlicz functions Φ,

$$C_n\left(d_F(\Phi)\right)^{\delta(\alpha)} \le \zeta_F(\Phi) \le D_n \max\left\{d_F(\Phi)^{\frac{1}{2\alpha}}, 1\right\}. \tag{164}$$

Proof: Let Φ be an α-Orlicz function, and let $F \subset \mathbf{R}^n$ such that $d_F(\Phi) < \infty$.. First we assume $\lim_{x\to\infty} \Phi(x)/x = \infty$.. (That is, we exclude the case $\Phi(x) = x$ for $x \ge 0$.) Then, by Lemma (5.2.10) M_Φ (defined in (130)) satisfies $M'_\Phi(x) > 0$ and $M''_\Phi(x) > 0$ for $x > 0$. (The case $\Phi(x)x, x \in [0,\infty)$, will be discussed later.). For M_Φ simplicity, we denote M_Φ by M. In (159), for each $\mu > 0$, the maximum of $\mu(x) - M(x)$.occurs at the unique point x satisfying $M'(x) = \mu$. Hence we can treat x as a function of μ, and write as a function satisfying the two equations

$M^*(\varphi) = ux - M(x)$,, where x is such that $M'(x) = u$.
(165)

We define M_2 on $[0,\infty)$ in a similar way by

$$M_2(w) = \sqrt{wx} - M(x), \text{ where } x \text{ is such that } M'(x) = \sqrt{w}. \tag{166}$$

Then for x satisfying (166),

$$M_2^1(w)=\frac{1}{2\sqrt{w}}x+\left(\sqrt{w}M^1(x)\right)\frac{dx}{dw}=\frac{x}{2M^1(x)} \tag{167}$$

and

$$M_2''(w)=\frac{M'(x)-xM''(x)}{2\left(M'(x)\right)^2}\frac{dx}{dw}. \tag{168}$$

By (167), (168) and Lemma (5.2.11) M_2 is an Orlicz function such that $M_2'(w)>0, M_2''(w)>0$ for $w>0$. By (159),

$$M_2^*(y)=\max\left\{yw-M_2(w):w>0\right\}. \tag{169}$$

For each $y>0$, the maximum of $yw-M_2(w)$ occurs at the unique point w satisfying $y=M_2'(w)$. Hence we can treat w as a function of y. But x is a function of w in (166). Therefore by (166) and (169),

$M_2^*(y)=-\frac{x}{2}M'(x)+M(x),$, here x is such that

$$\frac{x}{2M'(x)}=y. \tag{170}$$

Our aim is to apply (170) and the duality expressed in (160) to prove (164). To this end, let $s\in \mathrm{N}$, and consider a s-hypercube $A_1\times\ldots\times A_n\subset \mathbf{R}^n$ such that $\left|F\cap\left(A_1\times\ldots\times A_n\right)\right|\Psi_F$, By(158)

$$\left\|1_{F\cap\left(A_1\times\ldots\times A_n\right)}\right\|_{M_2^*}=\inf\left\{\rho>0:\sum_{W\in A_1\times\ldots\times A_n}M_2^*\left(1_F(w)/\rho\right)\le 1\right\}=\inf\left\{\rho>0:M_2^*(1/\rho)\Psi_F(s)\le 1\right\}. \tag{171}$$

Let $\rho_s>0$ be such that

$$M_2^*(1/\rho_s)\Psi_F(s)=1. \tag{172}$$

Then $\rho_s=\left\|1_{F\cap\left(A_1\times\ldots\times A_n\right)}\right\|_{M_2^*}$. Replacing y by $1/\rho_s$ in (170), and then combining (170) with (172), we have the system of equations

$$\frac{1}{\Psi_F(s)}=M(x)-\frac{x}{2}M'(x) \text{ and} \tag{173}$$

$$\rho_s=\frac{2M'(x)}{x}. \tag{174}$$

We want to estimate ρ_s using Eqs. (173) and (174). To this end, we first estimate x as a solution to Eq. (172). By (173),

$$M(x) \geq M(x) - \frac{x}{2}M'(x) = \frac{1}{\Psi_F(s)} \geq \frac{1}{d_F(\Phi)s^{\alpha}\phi(s)} \quad (by\ (75))$$

$$= \frac{1}{d_F(\Phi)s^{\alpha}\phi(s)}\left(\phi(s)\theta(1/\Theta(s))\right)^{\frac{1}{\alpha+1}} \quad (by\ (130))$$

$$= \frac{1}{d_F(\Phi)}\left(s^{-\frac{\alpha+1}{2}}\phi(s)^{-\frac{1}{2}}\right)^{\frac{2\alpha}{\alpha+1}}\theta(1/\Theta(s))^{\frac{1}{\alpha+1}}$$

$$= \frac{1}{d_F(\Phi)}(1/\Theta(s))^{\frac{2\alpha}{\alpha+1}}\left(\theta(1/\Theta(s))\right)^{\frac{1}{\alpha+1}} \qquad \text{by (127)} \qquad (175)$$

$$= \frac{1}{d_F(\Phi)}M(1/\Theta(s)) \quad (by\ (130))$$

if $d_F(\Phi) \geq 1$, then

$$M\left(d_F(\Phi)^{\frac{\alpha+1}{2\alpha}}x\right) = \left(\left(d_F(\Phi)\right)^{\frac{\alpha+1}{2\alpha}}x\right)^{\frac{2\alpha}{\alpha+1}}\left(\left(\theta\left(d_F(\Phi)\right)^{\frac{\alpha+1}{2\alpha}}x\right)\right)^{\frac{1}{\alpha+1}} \qquad \text{by (130)}$$

$$\geq d_F(\Phi)x^{\frac{2\alpha}{\alpha+1}}\left(\theta(x)\right)^{\frac{1}{\alpha+1}}(\text{because } \theta \text{ increasing})$$

$$d_F(\Phi)M(x) \quad (by\ (130))$$

$$\geq M(1/\Theta(s)) \qquad \text{by (175)} \qquad (176)$$

Because M is increasing, the comparison of both sides of (176) implies

$$x \geq \left(d_F(\Phi)\right)^{-\frac{\alpha+1}{2\alpha}}/\Theta(s) \qquad (177)$$

If $d_F(\Phi) < 1$, then by (175),

$$M(x) \geq M(1/\Theta(s)). \qquad (178)$$

Hence

$$x \geq 1/\Theta(s). \qquad (179)$$

For simplicity, let $\tilde{d}_F(\Phi) = \max\left(d_F(\Phi),1\right)$, By (177), (179),

$$x \geq \tilde{d}_F(\Phi)^{-\frac{\alpha+1}{2\alpha}} / \Theta(s), \tag{180}$$

which is the estimate that we need.

Now we estimate ρ_s . By (138), we have $0 \leq \frac{\theta'(x)}{\theta(x)} x \leq 1$ for all $x \geq 0$. Then by (131) and (174),

$$\rho_s = \frac{2}{\alpha+1} x^{\frac{2\alpha}{\alpha+1}-2} \left(\theta(x)\right)^{\frac{1}{\alpha+1}} \left\{2\alpha + \frac{\theta'(x)}{\theta(x)} x\right\} \leq 4\left(x^{-2}\theta(x)\right)^{\frac{1}{\alpha+1}}. \tag{181}$$

Because $0 \leq \frac{\theta'(x)}{\theta(x)} x \leq 1$ for all $x \geq 0, \left(x^{-2}\theta(x)\right)^{\frac{1}{\alpha+1}}$ is decreasing with increasing x. Then by (180) and (181),

$$\rho_s \leq 4\left(\left(\tilde{d}_F(\Phi)\right)^{-\frac{\alpha+1}{2\alpha}} / \Theta(s)\right)^{-\frac{2}{\alpha+1}} \left(\theta\left(\left(\tilde{d}_F(\Phi)\right)^{-\frac{\alpha+1}{2\alpha}} / \Theta(s)\right)\right)^{\frac{1}{\alpha+1}}$$

$$\leq 4\left(\tilde{d}_F(\Phi)\right)^{-\frac{1}{\alpha}} s\left(\phi(s)\right)^{\frac{1}{\alpha+1}} \left(\theta\left(1/\Theta(s)\right)\right)^{\frac{1}{\alpha+1}} \quad \text{(by (127) and because } \tilde{d}_F(\Phi) \geq 1)$$

$$= 4\left(\tilde{d}_F(\Phi)\right)^{\frac{1}{\alpha}} s \quad \text{(By (129))} \tag{182}$$

which is the estimate we need.

Let h be a function with support in f such that

$$\sum_{w \in A_1 \times \ldots \times A_n} M^*\left(\left|h(w)\right|\right) \leq 1.. \tag{183}$$

(165) and (166), for all $w \in A_1 \times \ldots \times A_n$,

$$M^*\left(\left|h(w)\right|\right) = M_2\left(\left|h(w)\right|^2\right). \tag{184}$$

Hence

$$\sum_{w \in A_1 \times \ldots \times A_n} M_2\left(\left|h(w)\right|^2\right) \leq 1. \tag{185}$$

Then,

$$\sum_{w\in A_1\times\ldots\times A_n}\left|h(w)\right|^2=\sum_{w\in A_1\times\ldots\times A_n}\left|h(w)\right|^2 1_F\le\left\|1_{F\cap(A_1\times\ldots\times A_n)}\right\|_{M_2^*}$$

by (185) and the duality in (160)

$$\le 2\left\|1_{F\cap(A_1\times\ldots\times A_n)}\right\|_{M_2^*}\quad\text{by (161)}$$

$$\le 8\left(\tilde{d}_F(\Phi)\right)^{1/\alpha}s\quad\text{by (182)}\qquad(186)$$

Let π_i be the canonical projection from $\mathbf{R}^n$ onto $\mathbf{R}$. By [219], there exists a cover $G_1,\ldots,G_n$ of $A_1\times\ldots\times A_n$ such that for every $i=1,\ldots,n,$,

$$\max_{r\in A_i}\sum_{w\in\pi_1^{-1}[r]}\left|h(w)\right|^2 1_{G_i}(w)\le 8\left(\tilde{d}_F(\Phi)\right)^{1/\alpha}.\qquad(187)$$

Suppose f is an R^n-polynomial in $C(\Omega^n)$ with spectrum in $A_1\times\ldots\times A_n$. (We identify $\left(r_{j_1},\ldots,r_{j_n}\right)\in\mathbf{R}^n$ with the character $w=r_{j_1}\otimes\ldots\otimes r_{j_n}$ on Ω^n.) By the Cauchy–Schwarz inequality, (187) and ([219]) we obtain for $i\in[n]$,

$$\left|\sum_{w\in A_1\times\ldots\times A_n}\hat{f}(w)h(w)1_G(w)\right|\le\sum_{r\in A_i}\left|\sum_{w\in\pi_1^{-1}[r]}\hat{f}(w)h(w)1_G(w)\right|$$

$$\le\sum_{r\in A_i}\left(\sum_{w\in\pi_1^{-1}[r]}\left|h(w)\right|^2 1_{G_i}(w)\right)^{1/2}\cdot\left(\sum_{w\in\pi_1^{-1}[r]}\left|\hat{f}(w)\right|^2\right)^{1/2}$$

$$\le\max_{r\in A_i}\left(\sum_{w\in\pi_1^{-1}[r]}\left|h(w)\right|^2 1_{G_i}(w)\right)^{1/2}\cdot\sum_{r\in A_i}\left(\sum_{w\in\pi_1^{-1}[r]}\left|\hat{f}(w)\right|^2\right)^{1/2}$$

$$\le 2\sqrt{2}\left(\tilde{d}_F(\Phi)\right)^{\frac{1}{2\alpha}}\zeta R(1)2^{\frac{n-1}{2}}\left\|f\right\|_{\infty,}\qquad(188)$$

where $\zeta_R(1)=\sup\left\|\hat{f}\right\|_{\ell^1}:f\in B_{C_R(\Omega)}\le 2.$ Therefore,

$$\left|\sum_{w\in A_1\times\ldots\times A_n}\hat{f}(w)h(w)\right|\le\sum_{i=1}^{n}\left|\sum_{w\in A_1\times\ldots\times A_n}\hat{f}(w)h(w)1_G(w)\right|$$

$$\leq 2\sqrt{2}n\left(\tilde{d}_F(\Phi)\right)^{\frac{1}{2\alpha}} \zeta R(1) 2^{\frac{n-1}{2}} \|f\|_\infty . \tag{189}$$

By (183), (189) and the duality in (160),

$$\|\hat{f}\|_M = \sup\left\{ \sum_{w\in A_1\times\ldots\times A_n} \hat{f}(w)h(w) : \sum_{w\in A_1\times\ldots\times A_n} M^*\left(|h(w)|\right) \leq 1 \right\}$$

$$\leq 2\sqrt{2}n\left(\tilde{d}_F(\Phi)\right)^{\frac{1}{2\alpha}} \zeta R(1) 2^{\frac{n-1}{2}} \|f\|_\infty \tag{190}$$

Then by (161),

$$\|\hat{f}\|_M \leq 2\sqrt{2}n\left(\tilde{d}_F(\Phi)\right)^{\frac{1}{2\alpha}} \zeta R(1) 2^{\frac{n-1}{2}} \|f\|_\infty . \tag{191}$$

which implies (164) with $D_n = 2\sqrt{2}n\zeta R(1) 2^{\frac{n-1}{2}}$.

Next suppose that $\Phi(x) = x$ for all $x \geq 0$. (Recall we excluded this case in the beginning of our proof.) Then $M_\Phi(x) = M(x) = x, x \geq 0$, and $\|\cdot\|_M = \|\cdot\|_{\ell^1(\mathbf{R}^n)}$.

Let h be in the unit ball of $\ell^\infty(\mathbf{R}^n)$ with support in F. Then

$$\sum_{w\in A_1\times\ldots\times A_n} |h(w)|^2 \leq |F \cap (A_1\times\ldots\times A_n)| \leq d_F(\Phi)s. \tag{192}$$

which corresponds to (186). Following the steps from (187) to (191), we have

$$\|\hat{f}\|_M \leq n\left(d_F(\Phi)\right)^{\frac{1}{2}} \zeta R(1) 2^{\frac{n-1}{2}} \|f\|_\infty . \tag{193}$$

which implies (164) in this case.

Now we prove the left side inequality of (164). For $s \in \mathrm{N}$, let $A_1 \times \ldots \times A_n$ be a s-hypercube in R^n such that $|F \cap (A_1 \times \ldots \times A_n)| = \Psi_F(s)$ Identify $(r_{j_1}, \ldots, r_{j_n}) \in \mathbf{R}^n$ with the character $w = r_{j_1} \otimes \ldots \otimes r_{j_n}$ on Ω^n By the Kahane–Salem–Zygmund probabilistic estimates ([219], Theorem X.8), there exists a $\{-1,+1\}$ – valued n-array $\{\in_w : \in_w = \pm, w \in F \cap (A_1 \times \ldots \times A_n)\}$ such that if

$$f_s = \frac{1}{s^{\frac{1}{2}}\left(\Psi_F(s)\right)^{\frac{1}{2}}} \sum_{w \in F \cap (A_1 \times \ldots \times A_n)} \in_w w \tag{194}$$

then

$$\|f_s\|_\infty \le C\|f_s\|_2 \left(\log\left(2^{ns}\right)\right)^{\frac{1}{2}} = Cs^{-\frac{1}{2}}(ns)^{\frac{1}{2}}(\log 2)^{\frac{1}{2}} = Cn^{\frac{1}{2}}(\log 2)^{\frac{1}{2}} \tag{195}$$

where $C > 0$ is a constant. By (158),

$$\left\|\hat{f}_s\right\|_{M\Phi} = \inf\left\{\rho > 0 \sum_{w \in F \cap (A_1 \times \ldots \times A_n)} M_\Phi\left(\left|f_s(w)\right| / \rho\right) \le 1\right\}$$

$$= \inf\left\{\rho > 0 \sum_{w \in F \cap (A_1 \times \ldots \times A_n)} M_\Phi\left(s^{-\frac{1}{2}}\left(\psi_F(s)\right)^{-\frac{1}{2}} / \rho\right) \le 1\right\}$$

$$= \inf\left\{\rho > 0 : M_\Phi\left(s^{-\frac{1}{2}}\left(\psi_F(s)\right)^{-\frac{1}{2}} \rho^{-1}\right)\psi_F(s) \le 1\right\}. \tag{196}$$

For each $s \in \mathrm{N}$, let $\rho_s > 0$ be such that

$$M_\Phi\left(s^{-\frac{1}{2}}\left(\Psi_F(s)\right)^{-\frac{1}{2}} / \rho_s^{-1}\right)\Psi_F(s) = 1. \tag{197}$$

Then $\rho_s = \left\|\hat{f}_s\right\|_{M_\Phi}$. By the definition of $d_F(\Phi)$ in (75)and because Φ is an α-Orlicz function, we have,

$$\Psi_F(s) \le d_F(\Phi)\Phi(s) = d_F(\Phi)s^\alpha \phi(s). \tag{198}$$

By the definition of M_Φ in (130) and by (197),

$$1 = \left(s^{-\frac{1}{2}}\left(\Psi_F(s)\right)^{-\frac{1}{2}} \rho_s^{-1}\right)^{\frac{2\alpha}{\alpha+1}} \left(\theta\left(s^{-\frac{1}{2}}\left(\Psi_F(s)\right)^{-\frac{1}{2}} \rho_s^{-1}\right)\right)^{-\frac{1}{\alpha+1}} \Psi_F(s)$$

$$\ge \rho_s^{-\frac{2\alpha}{\alpha+1}} \left(\Psi_F(s)\right)^{-\frac{1}{\alpha+1}} s^{-\frac{\alpha}{\alpha+1}} \left(\theta\left(s^{-\frac{\alpha+1}{2}}\left(d_F(\Phi)\right)^{-\frac{1}{2}}\left(\phi(s)\right)^{-\frac{1}{2}} \rho_s^{-1}\right)\right)^{\frac{1}{\alpha+1}}$$

by (198), and because θ is increasing (199)

By (199) and (127),

$$\rho_s^{2\alpha}\left(\Psi_F(s)\right)s^{-\alpha}\theta\left(\left(d_F(\Phi)\right)^{-\frac{1}{2}}\rho_s^{-1}/\Theta(s)\right). \tag{200}$$

Let

$$c=\left(d_F(\Phi)\right)^{-\frac{1}{2}}\rho_s^{-1} \tag{201}$$

By (200),

$$\rho_s^{2\alpha}\geq\Psi_F(s)s^{-\alpha}\theta\left(c/\Theta(s)\right) \tag{202}$$

If c >1, by (202) and (129),

$$\rho_s^{2\alpha}\geq\Psi_F(s)s^{-\alpha}\theta\left(1/\Theta(s)\right)=\Psi_F(s)s^{-\alpha}\left(\phi(s)\right)^{-1}=\frac{\Psi_F(s)}{\Phi(s)}. \tag{203}$$

Then (by taking supremum)

$$\sup\left\{\left\|\hat{f}_s\right\|_{M_\Phi}:s\in\mathrm{N}\right\}=\sup\left\{\rho_s:\in\mathrm{N}\right\}\geq\sup\left\{\Psi_F(s)/\Phi(s):s\in\mathrm{N}\right\}=\left(d_F(\Phi)\right)^{\frac{1}{2\alpha}}. \tag{204}$$

If $c\leq 1$by (127) and because ϕ is increasing,

$$c/\Theta(s)=cs^{-\frac{\alpha+1}{2}}\left(\phi(s)\right)^{-\frac{1}{2}}\geq\left(c^{-\frac{2}{\alpha+1}}s\right)^{-\frac{\alpha+1}{2}}\left(\phi\left(c^{\frac{2}{\alpha+1}}s\right)\right)^{-\frac{1}{2}}=1/\Theta\left(c^{-\frac{2}{\alpha+1}}s\right). \tag{205}$$

Then

$$\theta\left(c/\Theta(s)\right)\geq\theta\left(1/\Theta\left(1/\Theta\left(c^{-\frac{2}{\alpha+1}}s\right)\right)=\left(\phi\left(c^{-\frac{2}{\alpha+1}}s\right)\right)^{-1}\right)\text{By (129).} \tag{206}$$

By (202) and (206)

$$\rho_s^{2\alpha}\geq\Psi_F(s)s^{-\alpha}\left(\phi\left(c^{-\frac{2}{\alpha+1}}s\right)\right)^{-1}. \tag{207}$$

Because ϕ is concave and $c<1$,

$$\frac{\phi\left(c^{-\frac{2}{\alpha+1}}s\right)-\phi(0)}{c^{-\frac{2}{\alpha+1}}s}\leq\frac{\phi(s)-\phi(0)}{s} \tag{208}$$

Because $\phi(0)\ge 0$,

$$\frac{\phi\left(c^{-\frac{2}{\alpha+1}}s\right)}{\frac{2}{\alpha+1}}\le\phi(s)-\phi(0)+\frac{\phi(0)}{c^{-\frac{2}{\alpha+1}}}\le\phi(s). \tag{209}$$

By (207), (209) and (201),

$$\rho_s^{2\alpha}\ge\Psi_F(s)s^{-\alpha}(\phi(s))^{-1}c^{\frac{2}{\alpha+1}}=\frac{\Psi_F(s)}{\Phi(s)}(d_F\Phi)^{-\frac{1}{\alpha+1}}\rho_2^{-\frac{1}{\alpha+1}}. \tag{210}$$

Hence

$$\rho_2^{\frac{2(\alpha^2+\alpha+1)}{\alpha+1}}\ge\frac{\Psi_F(s)}{\Phi(s)}(d_F\Phi)^{-\frac{1}{\alpha+1}}. \tag{211}$$

Then (by taking supremum)

$$\left\{\sup\left\{\left\|\hat f\right\|_{M_\Phi}:s\in N\right\}=\sup\{\rho_s:s\in N\}\ge(d_F(\Phi))^{\left(1-\frac{1}{\alpha+1}\right)\left(\frac{\alpha+1}{2(\alpha^2+\alpha+1)}\right)}=(d_F(\Phi))^{\frac{\alpha}{2(\alpha^2+\alpha+1)}} \tag{212}$$

By (195), (204) and (212), we obtain the left side inequality of (164) with $\left(Cn^{1/2}(\log 2)^{1/2}\right)^{-1}$

Corollary (5.2.14)[218]: For $n\in N, F\subset\mathbf{R}^n$, and α-Orlicz function Φ,

$$\overline{\lim_{s\to\infty}}\frac{\Psi_F(s)}{\Phi(s)}<\infty\Leftrightarrow\zeta_F(\Phi)<\infty. \tag{213}$$

Remarks. ((5.2.15) [218]: (i) (A question) we were unable to answer the following: on the left side inequality in (164), can $\delta(\alpha)$ be replaced by $1/(2\alpha)$?

(ii) (Example(5.2.2). Let $\log^{(i)}$ denote the i-fold iteration of log. Suppose $\Phi(x)=x^\alpha\phi(x)$ is an α-Orlicz function such that for some $N>0$,

$$\phi(x)=\prod_{i=1}^{k}\left(\log^{(i)}x\right)^{\beta_i},\quad x\ge N, \tag{214}$$

for $k\ge 1$ and $\beta_i\ge 0$ for $i=1,\dots,k$, We want to show that the Orlicz function M_Φ defined in (130) can be approximated in a neighborhood of 0 by

$$M_{\alpha,\beta_1,\dots,\beta_k}(x)=\left(\frac{\alpha+1}{2}\right)^{\frac{\beta_1}{\alpha+1}}x^{\frac{2\alpha}{\alpha+1}}\prod_{i=1}^{k}\left(\log^{(i)}\frac{1}{x}\right)^{-\frac{\beta_i}{\alpha+1}}, \tag{215}$$

in the sense that $\lim_{x\to 0} M_{\alpha,\beta_1}(x)/M_\Phi(x)=1$. By (130) and (215),

$$\lim_{x\to 0}\frac{M_{\alpha,\beta_1,\ldots,\beta_k}(x)}{M_\Phi(x)}=\lim_{x\to 0}\frac{\left(\frac{\alpha+1}{2}\right)^{\frac{\beta_1}{\alpha+1}} x^{\frac{2\alpha}{\alpha+1}}\prod_{i=1}^k\left(\log^{(i)}\frac{1}{x}\right)^{-\frac{\beta_1}{\alpha+1}}}{x^{\frac{2\alpha}{\alpha+1}}(\theta(x))^{\frac{1}{\alpha+1}}}$$

$$=\left(\frac{\alpha+1}{2}\right)^{\frac{\beta_1}{\alpha+1}}\lim_{y\to\infty}\frac{\prod_{i=1}^k\left(\log^{(i)}\Theta(y)\right)^{-\frac{\beta_i}{\alpha+1}}}{(\theta(1/\Theta(y)))^{\frac{1}{\alpha+1}}} \text{ by substituting } x = 1/\Theta(y)$$

$$=\left(\frac{\alpha+1}{2}\right)^{\frac{\beta_1}{\alpha+1}}\lim_{y\to\infty}\phi(y)^{\frac{1}{\alpha+1}}\prod_{i=1}^k\left(\log^{(i)}\left(y^{\frac{\alpha+1}{2}}\left(\phi(y)\right)^{1/2}\right)\right)^{-\frac{\beta_i}{\alpha+1}} \text{ by (127) and (129)}$$

$$=\left(\frac{\alpha+1}{2}\right)^{\frac{\beta_1}{\alpha+1}}\lim_{y\to\infty}\prod_{i=1}^k\left(\log^{(i)}y\right)^{\frac{\beta_i}{\alpha+1}}\left(\frac{\alpha+1}{2}\log y\right)^{-\frac{\beta_i}{\alpha+1}}$$

$$\prod_{i=2}^k\left\{\log^{(i-1)}\left(\frac{\alpha+1}{2}\log y+\frac{1}{2}\log\left(\prod_{i=1}^k\left(\log^{(i)}x\right)^{\beta_i}\right)\right)\right\}^{-\frac{\beta_1}{\alpha+1}} \text{ by (124)}$$

$$=\prod_{i=2}^k\lim_{y\to\infty}\frac{\prod_{i=2}^k\left(\log^{(i)}y\right)^{\frac{\beta_i}{\alpha+1}}}{\prod_{i=2}^k\left\{\log^{(i-1)}\left(\frac{\alpha+1}{2}\log y+\frac{1}{2}\log\left(\prod_{i=1}^k\left(\log^{(i)}x\right)^{\beta_i}\right)\right)\right\}^{\frac{\beta_i}{\alpha+1}}}=1$$

(by L'Hopital's rule), (216)

as desired.

Definition (5.2.16)[218]:. (Cf. [219]).) For $n\in N, F\subset \mathbf{R}^n$, and α -Orlicz function Φ, let

$$\eta_F(\Phi)=\sup\left\{\|f\|_{L_p}/\Phi(p): p>2,\quad f\in B_{L_F^2(\Omega^\alpha)}\right\}. \tag{217}$$

This definition extends the definition in (69). Our aim is to establish a link between $\eta_F(\Phi)$and $d_F(\Phi)$, where $F\subset\mathbf{R}^n$. To this end, we first analyze analogous measurements in the context of $\left(\mathrm{T}^N\right)^n$, where $\mathrm{T}=\left\{e^{2\pi it}:t\in[0,1]\right\}$. We let $S=\{\beta_j: j\in N\}$ be the set of the canonicalprojections from T^N onto T:

$$\beta_j(t)=t(j),\quad t=(t(j): j\in N)\in\mathrm{T}^N. \tag{218}$$

We refer to S as the Steinhaus system, and view it as an independent set of characters on the compact Abelian group T^N with the normalized Haar measure P .

For $F \subset S^n$ and α-Orlicz function Φ, the definition of $d_F(\Phi)$ is the same as in (217). (Replace Ω^n by $\left(\mathrm{T}^N\right)^n$.)

Lemma (5.2.17)[218]: (Cf. [219]). For $n \in N, F \subset S^n$, and α-Orlicz function Φ,

$$16^{-n}\left(d_F(\Phi)\right)^{1/2} \leq \eta_F\left(\Phi^{1/2}\right) \leq \left(d_F(\Phi)\right)^{1/2}. \tag{219}$$

Proof: By [219], for all $f \in L_F^2\left(\left(T^N\right)^n\right)$,

$$\|f\|_{L^{2s}} \leq \left(\Psi_F(s)\right)^{1/2}\|f\|_{L^2}, \quad s \in N. \tag{220}$$

Because $\Psi_F(s) \geq d_F(\Phi)\Phi(s)$ for all $s \in N$, ,

$$\|f\|_{L^{2s}} \leq \left(d_F(\Phi)\right)^{1/2}\left(\Phi(s)\right)^{1/2}\|f\|_{L^2} \quad s \in N. \tag{221}$$

Let $\lambda = \dfrac{2s^2,2s...ps}{p}$. By Hölder's inequality, for $2s < p \leq 2s+2$,

$$\|f\|_{L^p} \leq \|f\|_{L^{5s}}^{\lambda}\|f\|_{L^{2s+2}}^{1-\lambda}. \tag{222}$$

Then by (221) and (222),

$$\|f\|_{L^p} \leq \left(\left(d_F(\Phi)\right)^{1/2}\left(\Phi(s)\right)^{1/2}\|f\|_{L^2}\right)^{\lambda}\left(\left(d_F(\Phi)\right)^{1/2}\left(\Phi(s+1)\right)^{1/2}\|f\|_{L^2}\right)^{1-\lambda} \tag{223}$$

$$= \left(d_F(\Phi)\right)^{1/2}\left(\left(\Phi(s)\right)^{\lambda}\left(\Phi(s+1)\right)^{1-\lambda}\right)^{1/2}\|f\|_{L^2}.$$

Because

$$p > 2s \geq s+1 \tag{224}$$

and Φ is increasing,

Therefore,

$$\|f\|_{L^p} \leq \left(d_F(\Phi)\right)^{\frac{1}{2}}\left(\Phi(P)\right)^{\frac{1}{2}}\|f\|_{L^2}. \tag{225}$$

To verify the left side inequality of (219), let $s \in N$ and let $A_1 \times \ldots \times A_n$ be a s-hypercube in S^n such that $\left|F \cap (A_1 \times \ldots \times A_n)\right| = \Psi_F(s)$. Consider the Riesz product

$$H_s = \prod_{\beta \in A_1}\left(1 + \frac{\beta + \overline{\beta}}{2}\right) \otimes \ldots \otimes \prod_{\beta \in A_n}\left(1 + \frac{\beta + \overline{\beta}}{2}\right). . \tag{226}$$

Then $\|H_s\|_{L^1} = 1$ and $\|H_s\|_{L^2} = 2^{ns/2}$. Hence for $1 \le p \le 2$, ,

$$\|H_s\|_{L^p} \le \|H_s\|_{L^1}^{1-\frac{2}{q}} \|H_s\|_{L^2}^{\frac{2}{q}} = 2^{\frac{ns}{q}}, \qquad 1/p + 1/q = 1.. \tag{227}$$

Let

$$h_s = \sum_{(\beta_1, \ldots, \beta_n) \in F \cap (A_1 \times \ldots \times A_n)} \beta_1 \otimes \ldots \otimes \beta_n. \tag{228}$$

Let E (expectation) denote integration with respect to Haar measure, either on Ω or on T^N. Let E^n denote the n-fold iteration of E. By Hölder's inequality and (227) with $q = x$,

$$\left|\mathrm{E}^n H_s h_s\right| \le \|H_s\|_{L^p} \|h_s\|_{L^q} \le 2^{\frac{ns}{s}} \|h_s\|_{L^s} \le 2^n \eta_F(\Phi)\Phi(s)\|h_s\|_{L^2}. \tag{229}$$

Because

$$\left|\mathrm{E}^n H_s h_s\right| = 2^{-n}\Psi_F(s), \tag{230}$$

and

$$\|h_s\|_{L^2} = \left(\Psi_F(s)\right)^{\frac{1}{2}}, \tag{231}$$

we obtain

$$4^{-n}\left(\Psi_F(s)\right)^{\frac{1}{2}} \le \eta_F(\Phi)\Phi(s), \tag{232}$$

which implies the left side of (219).

Corollary (5.2.18) [218]: (Cf. [219]). For $n \in N, F \subset \mathbf{R}^n$, and α-Orlicz function Φ,

$$16^{-n}\left(d_F(\Phi)\right)^{\frac{1}{2}} \le \eta_F\left(\Phi^{\frac{1}{2}}\right) \le 4^n\left(d_F(\Phi)\right)^{\frac{1}{2}}. \tag{233}$$

Proof. For each $j \in N$, let r_j be the Rademacher function in R such that

$$r_j(w) = w(j), \quad w \in \Omega = \{-1,1\}^N, \tag{234}$$

and let β_j be the Steinhaus function in S such that

$$\beta_j(t) = t(j), \quad t \in T^N. \tag{235}$$

Let f be an F-polynomial (i.e., spect f = support $\hat{f} \subset F$,, and spect f is finite). Define for $t = (i_1,...,i_n) \in \left(T^N\right)^n$,

$$f_t = \sum_{(r_{j_1},...,r_{j_n}) \in F} \hat{f}\left(r_{j_1} \otimes ... \otimes r_{j_n}\right)\beta_{j_1}(t_1)...\beta_{j_n}(t_n) r_{j_1} \otimes ... \otimes r_{j_n}. \tag{236}$$

For $t \in \left(T^N\right)^n$.there exists $\theta_t \in L^1(\Omega^n)$ such that

$$\hat{\theta}_t\left(r_{j_1} \otimes ... \otimes r_{j_n}\right) = \overline{\beta_{j_1}(t_1)}...\overline{\beta_{j_n}(t_n)}, \qquad r_{j_1} \otimes ... \otimes r_{j_n} \in spect\, f, \tag{237}$$

$$\|\theta_t\|_{L^1} \le 4^n \tag{238}$$

(See [219]). Then

$$\|f\|_{L^q}^q = \|f_t * \theta_t\|_{L^q}^q \le 4^{nq}\|f_t\|_{L^q}^q, \tag{239}$$

where⋆denotes convolution. Integrating both sides of (239) with respect to the Haar measure on $\left(T^N\right)^n$, applying Fubini's Theorem, and then the right side of (219), we obtain

$$\|f\|_{L^q} \le 4^n\left(d_F(\Phi)\right)^{\frac{1}{2}}\left(\Phi(q)\right)^{\frac{1}{2}}\|f\|_{L^2}, \tag{240}$$

which implies the right side of (173).The proof of the left side of (233) is a transcription of the proof of the left side of (219).

Suppose $(A,)$is a probability space. For any Orlicz function Ψ, consider the Orlicz norm corresponding to Ψ,

$$\|X\|_\psi = \inf\left\{\rho > 0 : E\psi\left(|X|/\rho\right) \le 1\right\}, \qquad X \in L^0(\mathscr{A}, \mathscr{P}). \tag{241}$$

The classical Kahane–Khintchin inequality states that: if $\Psi(x) = \exp\left(x^2\right) - 1$ for $x \ge 0$,, then there exists $K > 0$ such that,

$$\|X\|_{\psi} \leq K\|X\|_{L^2(\Omega,p)}, \qquad X \in L_R^2(\Omega, P). \tag{242}$$

(see [87].)We will extend the inequality in (183) to $F \subset \mathbf{R}^n$ Let $\Phi = x^{\alpha}\phi(x)$ be an α-Orlicz

function (as per Definition (5.2.1). Define

$$f(x) = x^{\alpha}\left(\phi\left(x^2\right)\right)^{\frac{1}{2}}, \qquad x \geq 0, \tag{243}$$

and let

$$g = f^{-1}. \tag{244}$$

Lemma (5.2.19)[218]: (Cf. [219].) Suppose (A,P) is a probability space, and $X \in B_{L^2(A)}$.

Then the following are equivalent:

(i) there exists $0 < A < \infty$ such that

$$\overline{\lim_{x\to\infty}} \exp\left(A\left(g(x)\right)^2\right)\mathscr{P}\left(|X| > x\right) < \infty; \tag{245}$$

(ii) there exists $0 < B < \infty$ such that

$$\sup\left\{\|X\|_{L^p} / p^{\frac{\alpha}{2}}\left(\phi(p)\right)^{\frac{1}{2}} : p > 2\right\} \leq B; \tag{246}$$

(iii) there exists $0 < C < \infty$ such that

$$\overline{\lim_{x\to\infty}} \mathrm{E}\exp\left(tg\left(|X|\right) - Ct^2\right) < \infty; \tag{247}$$

(iv) there exist $0 < D < \infty$ such that

$$\mathrm{E}\exp\left(D\left(g\left(|X|\right)\right)^2\right) < \infty. \tag{248}$$

Proof: (i) $\Rightarrow$(ii). Suppose $\overline{\lim_{x\to\infty}} \exp\left(A\left(g(x)\right)^2\right)\mathscr{P}\left(|X| > x\right) := B_1 < \infty$, For $p > 2$ sufficiently large,

$$\mathrm{E}|\mathrm{X}|^p = \int_0^{\infty} \mathscr{P}\left(|X|^p > x\right) dx$$

$$\le p^{\frac{\alpha p}{2}}\left(\phi(p)\right)^{\frac{p}{2}}+B_1\int_{p^{\frac{\alpha p}{2}}(\phi(p))^{\frac{p}{2}}}^{\infty}\exp\left(-A\left(g\left(x^{\frac{1}{p}}\right)\right)^2\right)dx \tag{249}$$

let $V=g\left(x^{\frac{1}{p}}\right)$. Then

$$x=\left(g^{-1}(y)\right)^p=\left(f(y)\right)^p=y^{\alpha p}\left(\phi\left(y^2\right)\right)^{\frac{p}{2}}. \tag{250}$$

Hence

$$dx/dy=\alpha p y^{\alpha p-1}\left(\phi\left(y^2\right)\right)^{\frac{p}{2}}\left\{1+\frac{1}{\alpha}\frac{\phi'\left(y^2\right)}{\phi\left(y^2\right)}y^2\right\}$$

$$\le 2\alpha p y^{\alpha p-1}\left(\phi\left(y^2\right)\right)^{\frac{p}{2}}.\ \text{(by(126))}. \tag{251}$$

When $x=P^{\frac{\alpha p}{2}}\left(\phi(p)\right)^{\frac{p}{2}}$, we have $y=\sqrt{p}.$ Hence by (249) and (251),

$$\mathrm{E}|X|^p\le p^{\frac{\alpha p}{2}}\left(\phi(p)\right)^{\frac{p}{2}}+B_1\int_{\sqrt{p}}^{\infty}2\alpha p y^{\alpha p-1}\left(\phi\left(y^2\right)\right)^{\frac{p}{2}}\exp\left(-Ay^2\right)dy. \tag{252}$$

By the Cauchy–Schwarz inequality,

$$\mathrm{E}|X|^p\le p^{\frac{\alpha p}{2}}\left(\phi(p)\right)^{\frac{p}{2}}+2B_1\alpha p\sqrt{\pi/A}\left\{\int_0^{\infty}\sqrt{\pi/A}\,y^{2(\alpha p-1)}\exp\left(-Ay^2\right)dy\right\}^{\frac{1}{2}}$$

$$\left\{\int_{\sqrt{p}}^{\infty}\sqrt{A/\pi}\left(\phi\left(y^2\right)\right)^p\exp\left(-Ay^2\right)dy\right\}^{\frac{1}{2}}. \tag{253}$$

The first integral on the right side of (253) is the $2(\alpha p-1)$ moment of a Gaussian random variable with mean 0 and variance $1/2A$. Hence there exists $B_2>0$ such that

$$\int_0^{\infty}\sqrt{A/\pi}\,y^{2(\alpha p-1)}\exp\left(-Ay^2\right)dy\le B_2^p\,p^{\alpha p-1}. \tag{254}$$

Next we estimate the second integral on the right side of (253). By property (iii) in Definition (5.2.1), $\phi\left(y^2\right)/y$ is eventually decreasing. Because p is sufficiently large, for all $y\ge\sqrt{p}$,

$$\phi(y^2)/y \le \phi(p)/\sqrt{p}. \tag{255}$$

Then

$$\frac{1}{(\phi(p))^p}\int_{\sqrt{p}}^{\infty}\sqrt{A/\pi}\left(\phi\left(y^2\right)\right)^p \exp\left(-Ay^2\right)dy$$

$$= \int_{\sqrt{p}}^{\infty}\sqrt{A/\pi}\left(\phi\left(y^2\right)/\phi(p)\right)^p \exp\left(-Ay^2\right)dy$$

$$\le \int_{\sqrt{p}}^{\infty}\sqrt{A/\pi}\left(y\sqrt{p}\right)^p \exp\left(-Ay^2\right)dy \qquad \text{(by (255))}$$

$$\le \frac{1}{p^{\frac{p}{2}}}\int_{0}^{\infty}\sqrt{A/\pi}\, y^p \exp\left(-Ay^2\right)dy \le B_3^B \tag{256}$$

for some $B_3 > 0$ (by estimating p-th moments of Gaussian random variables). Then

$$\int_{\sqrt{p}}^{\infty}\sqrt{A/\pi}\left(\phi\left(y^2\right)\right)^p \exp\left(-Ay^2\right)dy \le B_3^p\left(\phi(p)\right)^p. \tag{257}$$

By (253), (254) and (257), there exists $B > 0$ such that

$$\|X\|_{Lp} \le Bp^{\frac{\alpha}{2}}\left(\phi(p)\right)^{\frac{1}{2}}. \tag{258}$$

$(ii) \Rightarrow (iii)$.. We assume $B \ge 1$. For t > 0,

$$\mathrm{E}\exp\left(tg\left(|X|\right)\right) = \sum_{k=0}^{\infty}\frac{t^k}{k!}\mathrm{E}\left(tg\left(|X|\right)\right)^k. \tag{259}$$

For each $k \ge 1$, let

$$f_k(x) = x^{\alpha}\left(\phi\left(x^{\frac{2}{k}}\right)\right)^{\frac{k}{2}}, \qquad x \in [0,\infty), \tag{260}$$

and let

$$g_k = f_k^{\frac{1}{k}}. \tag{261}$$

Then $f_1 = f$ and $g_1 = g$. We will show that g_k is increasing for $k \ge 1$, and is concave for $k \ge 2$. To this end, it suffices to show that f_k is increasing for $k \ge 1$, and is convex for $k \ge 2$. By (260),

$$f_k'(x) = x^{\alpha-1}(\phi(y))^{\frac{1}{2}}\left\{\alpha + \frac{\phi'(y)}{\phi(y)}y\right\} \ge 0, \tag{262}$$

Where $y = x^{2/k}$. Hence $g_k = f_k^{-1}$ is increasing for all $k \ge 1$. By (262),

$$f_k''(x) = x^{\alpha-2}(\phi(y))^{\frac{1}{2}}\left(\frac{2}{k}\right)\left\{\frac{k\alpha(\alpha-1)}{2} + \left(k\alpha - \frac{k}{2}+1\right)\frac{\phi'(y)}{\phi(y)}y + \frac{\phi''(y)}{\phi(y)}y^2 + \left(\frac{k}{2}-1\right)\left(\frac{\phi'(y)}{\phi(y)}\right)^2 y\right\}. \tag{263}$$

Because Φ is an Orlicz function, for all $x \ge 0$,

$$\Phi''(x) = \left(x^\alpha\phi(x)\right)'' = x^{\alpha-2}\phi(x)\left\{\alpha(\alpha-1) + 2\alpha\frac{\phi'(x)}{\phi(x)}x + \frac{\phi''(x)}{\phi(x)}x^2\right\} \ge 0. \tag{264}$$

For $k \ge 2$, the expression inside the brackets of (263) is

$$\frac{k\alpha(\alpha-1)}{2} + \left(k\alpha - \frac{k}{2}+1\right)\frac{\phi'(y)}{\phi(y)}y + \frac{\phi''(y)}{\phi(y)}y^2 + \frac{k}{2}-1\left(\frac{\phi'(y)}{\phi(y)}y\right)^2$$

$$\ge \alpha(\alpha-1) + 2\alpha\frac{\phi'(y)}{\phi(y)}y + \frac{\phi''(y)}{\phi(y)}y^2$$

≥ 0 by (264) (265)

Hence $f_k'' \ge 0$ for $k \ge 2$. Therefore $g_k = f_k^{-1}$ is concave for $k \ge 2$. as desired.

By (243), (244), (260) and (261),

$$(g(x))^k = \left(f^{-1}(x)\right)^k = f_k^{-1}\left(x^k\right) = g_k\left(x^k\right). \tag{266}$$

Then, by Jensen's inequality, for $k = 2$,

$$\mathrm{E}\left(g(|X|)\right)^k = \mathrm{E}\left(g_k\left(|X|^k\right)\right) \le g_k\left(\mathrm{E}|X|^k\right). \tag{267}$$

By assumption (ii) and because $\|X\|_{L^2} \le 1$, for $k \ge 2$,

$$\mathrm{E}|X|^k \le B^k K^{\frac{\alpha k}{2}}(\phi(k))^{\frac{k}{2}}. \tag{268}$$

Because $B \ge 1$, we have $\phi\left(B^{2/\alpha}k\right) \ge \phi(k)$. Hence

$$\mathrm{E}|X|^k \le \left(B^{\frac{2}{\alpha}}k\right)^{\frac{\alpha k}{2}} \left(\phi\left(B^{\frac{2}{\alpha}}k\right)\right)^{\frac{k}{2}} = f_k\left(\left(B^{\frac{2}{\alpha}}k\right)^{\frac{k}{2}}\right). \tag{269}$$

By (267) and (269), for $k \ge 2$,

$$\mathrm{E}\left(g\left(|X|\right)\right)^k \le \left(g_k o f_k\right)\left(\left(B^{\frac{2}{\alpha}}k\right)^{\frac{k}{2}}\right) = \left(B^{\frac{2}{\alpha}}k\right)^{\frac{k}{2}} = B^{\frac{k}{\alpha}}k^{\frac{1}{2}} \ . \tag{270}$$

Next we estimate $\mathrm{E}\left(g\left(|X|\right)\right)$. By (243), $0 \le f(x) \le \left(\phi(1)\right)^{1/2}$ for $0 \le x \le 1$. Then
$0 \le g(x) \le 1$

for $0 \le x \le \left(\phi(1)\right)^{1/2}$ (because $\left(g = f^{-1}\right)$. Also by (243), for $x \ge \left(\phi(1)\right)^{1/2}$, we have

$$g(x) \le x^{1/\alpha}\left(\phi(1)\right)^{-\frac{1}{2\alpha}} = \left(x\left(\phi(1)\right)^{-\frac{1}{2}}\right)^{1/\alpha}$$

$$\le \left(x\left(\phi(1)\right)^{-\frac{1}{2}}\right)^2 \text{(because } x\left(\phi(1)\right)^{-\frac{1}{2}} \ge 1\text{)}$$

$$= x^2\left(\phi(1)\right)^{-1}. \tag{271}$$

Let $k = \max\left\{2\left(\phi((1))\right)^{-1}, 2\right\}$. Then

$$\mathrm{E}\left(g\left(|X|\right)\right) = \mathrm{E}\left(g\left(|X|\right)\right)1_{\left\{|X| \le (\phi(1))^{\frac{1}{2}}\right\}} \mathrm{E}\left(g\left(|X|\right)\right)1_{\left\{|X| > (\phi(1))^{\frac{1}{2}}\right\}}$$

$$\le 1 + \left(\phi(1)\right)^{-1}\mathrm{E}|X|^2 \le K \text{ (because } \mathrm{E}|X|^2 \le 1\text{)}. \tag{272}$$

Applying (270) and (272) to (259), we obtain for t sufficiently large,

$$\mathrm{E}\exp\left(tg\left(|X|\right)\right) \le 1 + Kt + \sum_{k=2}^{\infty}\frac{t^k}{K!}B^{\frac{k}{\alpha}}K^{\frac{k}{2}}$$

$$\le \exp\left(Ct^2\right) \quad \text{(because } K^{\frac{k}{2}}/K! < 2^k/(K/2)!\text{)} \tag{273}$$

for some $C > 0$.

$(iii) \Rightarrow (i)$. Because g is increasing $\left(g = g_1\right)$, for $x > 0$ and $t > 0$,

$\mathscr{P}\left(|X| > x\right) \le \mathscr{P}\left(g\left(|X|\right) > g(x)\right)$

$$\leq \frac{\mathrm{E}\exp\left(tg\left(|X|\right)\right)}{\exp\left(tg\left(x\right)\right)} \text{ (by Chebyshev's inequality).} \tag{274}$$

Then by assumption (iii), for $t > 0$ sufficiently large,

$$\mathscr{P}\left(|X| > x\right) \leq \frac{\exp\left(Ct^2\right)}{\exp\left(tg\left(x\right)\right)}. \tag{275}$$

Put $t = g(x)/2C$ in (274), and obtain (245) with $A = 1/4C$.

$(i) \Rightarrow (iV)$. Suppose $\overline{\lim}_{x\to\infty} \exp\left(A\left(g\left(x\right)\right)^2\right)\mathscr{P}\left(|X| > x\right) : M_1 < \infty$, and let $M_2 > 0$ be sufficiently large so that $\mathscr{P}\left(|X| > x\right) \leq M_1 \exp\left(-A\left(g\left(x\right)\right)^2\right)$ for $x \geq M_2$. Choose $0 < D < A$. Then

$$\mathrm{E}\exp\left(D\left(g\left(|X|\right)\right)^2\right) = \int_0^\infty \mathscr{P}\left(\exp\left(D\left(g\left(|X|\right)\right)^2\right) > x\right)dx$$

$$\leq M_2 + \int_{M_2}^\infty \mathscr{P}\left(|X| > g^{-1}\left(\log x\right)^{\frac{1}{2}} / D^{\frac{1}{2}}\right)dx \text{ (because g is increasing)}$$

$$\leq M_2 + M_1 \int_{M_2}^\infty \exp\left\{-A\left[g\left(g^{-1}\left(\left(\log x\right)^{\frac{1}{2}} / D^{\frac{1}{2}}\right)\right)\right]^2\right\}dx \text{ (by assumption} (i) \text{) (i)}$$

$$= M_2 + M_1 \int_{M_2}^\infty x^{-\frac{A}{D}}dx \leq M_2 + M_1. \tag{276}$$

$(iV) \Rightarrow (i)$. Because g is increasing, for $x > 0$ sufficiently large,

$$\mathscr{P}\left(|X| > x\right) \leq \mathscr{P}\left(D\left(g\left(|X|\right)\right)^2 > D\left(g\left(x\right)\right)^2\right)$$

$$\leq \frac{\mathrm{E}\exp\left(D\left(g\left(|X|\right)\right)^2\right)}{\exp\left(D\left(g\left(x\right)\right)^2\right)} \text{(by Chebyshev's inequality),} \tag{277}$$

which implies (245).

Lemma (5.2.20)[218]: Let f and g be the functions defined in (243) and (244). Let

$$h\left(x\right) = \exp\left(\left(g\left(x\right)\right)^2\right) - 1, \qquad x \geq 0. \tag{278}$$

Then there exists $N > 0$ such that $h''(x) \geq 0$ for all $x \geq N.$.

Proof:

$$h'(x) = 2\exp\left(\left(g(x)\right)^2\right)g(x)g'(x), \tag{279}$$

and

$$h''(x) = 2\exp\left(\left(g(x)\right)^2\right)I(x), \tag{280}$$

Where

$$I(x) = 2\left(g(x)\right)^2\left(g'(x)\right)^2 + \left(g'(x)\right)^2 + g(x)g''(x). \tag{281}$$

Because $g = f^{-1}$, we have

$$g'\left(f(x)\right)f'(x) = 1, \qquad x \geq 0. \tag{282}$$

Hence

$$g''\left(f(x)\right)\left(f'(x)\right)^2 + g'\left(f(x)\right)f''(x) = 0. \tag{283}$$

By (282) and (283),

$$g''\left(f(x)\right) = -\frac{f''(x)}{\left(f'(x)\right)^3}. \tag{284}$$

By (281) and (284)

$$I\left(f(x)\right) = 2\left(g\left(f(x)\right)\right)^2\left(g'\left(f(x)\right)\right)^2 + \left(g'\left(f(x)\right)\right)^2 + g\left(f(x)\right)g''\left(f(x)\right)$$

$$= 2x^2\left(g'\left(f(x)\right)\right)^2 + \left(g'\left(f(x)\right)\right)^2 + xg''\left(f(x)\right) \text{(because } g = f^{-1}\text{)}$$

$$= 2x^2\frac{1}{\left(f'(x)\right)^2} + \frac{1}{\left(f'(x)\right)^2} - x\frac{f''(x)}{\left(f'(x)^3\right)} \text{by (282) and (284)}$$

$$= \frac{1}{\left(f'(x)\right)^2}\left\{2x^2 + 1 - \frac{f''(x)}{f'(x)}x\right\}. \tag{285}$$

By (202) with $K = 1$,

$$f'(x) = x^{\alpha-1}\left(\phi\left(x^2\right)\right)^{\frac{1}{2}}\left\{\alpha + \frac{\phi'\left(x^2\right)}{\phi\left(x^2\right)}x^2\right\} \geq \alpha x^{\alpha-1}\left(\phi\left(x^2\right)\right)^{\frac{1}{2}}. \qquad (286)$$

By (263) with k = 1, and because $0 \leq \frac{\phi'\left(x^2\right)}{\phi\left(x^2\right)}x^2 \leq 1$ and $\phi'' \leq 0$,

$$f''(x) = x^{\alpha-2}\left(\phi\left(x^2\right)\right)^{\frac{1}{2}}\left\{\alpha(\alpha-1) + (2\alpha+1)\frac{\left(\phi'\left(x^2\right)\right)}{\phi\left(x^2\right)}x^2 + 2\frac{\left(\phi''\left(x^2\right)\right)}{\phi\left(x^2\right)}x^4 - \frac{\phi'\left(x^2\right)}{\phi\left(x^2\right)}x^2\right\}$$

$$\leq x^{\alpha-2}\left(\phi\left(x^2\right)\right)^{\frac{1}{2}}\left\{\alpha(\alpha-1) + 2\alpha + 1\right\}. \qquad (287)$$

Hence

$$\frac{f''(x)}{f(x)}x \leq \alpha + 1 + \frac{1}{\alpha} \leq \alpha + 2. \qquad (288)$$

By (285) and (288),

$$I\left(f(x)\right) \geq \frac{1}{\left(f'(x)\right)^2}\left\{2x^2 - \alpha - 1\right\}. \qquad (289)$$

Replacing x by $g(x)$ in (289), we have

$$I(x) \geq \frac{1}{\left(f'\left(g(x)\right)\right)^2}\left\{2\left(g(x)\right)^2 - \alpha - 1\right\}. \qquad (290)$$

Then for $x \geq g^{-1}\left((\alpha+1)/2\right)^{\frac{1}{2}}$, we have $I(x) \geq 0$ which implies $h''(x) \geq 0$.

Let Ψ_Φ be an Orlicz function such that for some $N > 0$,

$$\Psi_\Phi(x) = \exp\left(\left(g(x)\right)^2\right) - 1, \qquad x \geq N, \qquad (291)$$

where g is defined in (244).

Lemma (5.2.21)[218]: (Cf. [219]).Suppose $(\mathscr{A}, \mathscr{P})$ is a probability space, and

$X \in B_{L_F^2(\mathscr{A},\mathscr{P})}$.. Then the following are equivalent:

(i) there exists $0 < D < \infty$ such that

$$\mathrm{E}\exp\left(D\left(g\left(|X|\right)\right)^2\right) < \infty; \tag{292}$$

$$\|X\|_{\Psi_\Phi} < \infty. \tag{293}$$

Proof: $(i) \Rightarrow (ii)$. Suppose

$$\mathrm{E}\exp\left(D\left(g\left(|X|\right)\right)^2\right) \le M, \tag{294}$$

for some $M \ge 1$.. Let $\beta > 0$ be such that $\beta \ge \max\{4M, D\}$ and $\Psi_\Phi\left(< / \beta\right) \le \frac{1}{2}$.. Then

$$\mathrm{E}_{\Psi_\Phi}\left(|X| / \beta\right) 1_{\{|X| < N\}} \le \frac{1}{2}. \tag{295}$$

Because ϕ is concave, we have for $c \ge 1$ and $x > 0$,

$$\frac{\phi(cx) - \phi(0)}{cx} \le \frac{\phi(x) - \phi(0)}{x}. \tag{296}$$

Then, because $\phi(0) \ge 0$,,

$$\frac{\phi(cx)}{cx} \le \frac{\phi(x)}{x} - \frac{\phi(x)}{x} + \frac{\phi(0)}{cx} \le \frac{\phi(x)}{x}. \tag{297}$$

Let

$$L(x) = g\left(\left(\beta D^{-1}\right)^{-\frac{\alpha+1}{2}} x\right), \qquad x \ge 0. \tag{298}$$

Then

$$x = \left(\beta D^{-1}\right)^{\frac{\alpha+1}{2}} \left(\beta D^{-1}\right)^{\frac{\alpha+1}{2}} x = \left(\beta D^{-1}\right)^{\frac{\alpha+1}{2}} g^{-1}\left(L(x)\right)$$

$$= \left(\beta D^{-1}\right)^{\frac{\alpha+1}{2}} \left(L(x)\right)^{\alpha} \left(\phi\left(\left(L(x)\right)^2\right)\right)^{\frac{1}{2}} \text{ (by (244) and (245))}$$

$$=\left(\beta D^{-1}\right)^{\frac{\alpha}{2}}\left(L(x)\right)^{\alpha}\left(\beta D^{-1}\phi\left(\left(L(x)\right)^{2}\right)\right)^{\frac{1}{2}}$$

$$=\left(\beta D^{-1}\right)^{\frac{1}{2}}\left(L(x)\right)^{\alpha}\left(\beta D^{-1}\phi\left(\left(L(x)\right)^{2}\right)\right)^{\frac{1}{2}}. \text{ (by (297) and because } \beta D^{-1}\geq 1)$$

$$=g^{-1}\left(\left(\beta D^{-1}\right)L(x)\right) \qquad \text{by (243) and (244).} \tag{299}$$

By (298) and (299),

$$g(x)\geq\left(\beta D^{-1}\right)^{\frac{1}{2}}L(x)=\left(\beta D^{-2}\right)^{\frac{1}{2}}g\left(\left(\beta D^{-1}\right)^{-\frac{\alpha+1}{2}}x\right). \tag{300}$$

Because $\beta\geq 2M$, we have

$$(2M)^{-\frac{1}{2}}D^{\frac{1}{2}}g(x)\geq g\left(\left(\beta D^{-1}\right)^{-\frac{\alpha+1}{2}}x\right). \tag{301}$$

Then

$$\mathrm{E}\exp\left(g\left(\left(\beta D^{-1}\right)^{-\frac{\alpha+1}{2}}|X|\right)\right)^{2}\leq 1+\mathrm{E}\sum_{k=1}^{\infty}\frac{1}{k!}\left((2M)^{-1}D\left(g(|X|)\right)^{2}\right)^{k}$$

$$\leq 1+\mathrm{E}\sum_{k=1}^{\infty}\frac{1}{k!}\left(D\left(g(|X|)\right)^{2}\right)^{k} \text{ (because } M\geq 1)$$

$$\leq 1+\frac{1}{2M}\mathrm{E}\exp\left(D\left(g(|X|)\right)^{2}\right)$$

$$\leq\frac{3}{2} \qquad \text{by (294)} \tag{302}$$

By the definition of ψ_{Φ} in (291), we have

$$\mathrm{E}_{\Psi\Phi}\left(\left(\beta D^{-1}\right)^{-\frac{\alpha+1}{2}}|X|\right)1_{(|X|\geq N)}=\mathrm{E}\exp\left(g\left(\left(\beta D^{-1}\right)^{\frac{\alpha+1}{2}}|X|\right)\right)^{2}1_{(|X|\geq N)}-1$$

$$\leq\frac{1}{2}\text{by(302).} \tag{303}$$

Let $K=\max\left\{\beta,\left(\beta D^{-1}\right)^{\frac{\alpha+1}{2}}\right\}$. By (295) and (303), we have

$$\mathrm{E}\Psi_{\Phi}\left(|X|/K\right)\leq 1. \tag{304}$$

Therefore

$$\|X\|_{\Psi_\Phi} \le K. \tag{305}$$

$(iii)\Rightarrow(i)$. If $\|X\|_{\psi_\Phi} \le K$ for some $K > 0$, then

$$\mathrm{E}\Psi_\Phi\left(|X|/K\right) \le 1. \tag{306}$$

Hence by the definition of Ψ_Φ in (291),

$$\mathrm{E}\left\{\exp\left(\left(g\left(|X|/K\right)\right)^2\right)-1\right\}1_{\{|X|\ge N\}}\le 1. \tag{307}$$

Let

$$M = \max\left\{4, 2\mathrm{E}\exp\left(\left(g\left(N/K\right)\right)^2\right)\right\}. \tag{308}$$

By (307) and (308),

$$\mathrm{E}\exp\left(\left(g\left(|X|/K\right)\right)^2\right) \le \frac{M}{2}+2 \le M. \tag{309}$$

We may assume $K \ge 1$. By (303) and (304), for $x \ge 0$,

$$x = f\left(g\left(x\right)\right) = \left(g\left(x\right)\right)^\alpha \left(\phi\left(\left(g\left(x\right)\right)^2\right)\right)^{\frac{1}{2}}. \tag{310}$$

Then

$$\begin{aligned} f\left(g\left(x\right)/K^{\frac{1}{\alpha}}\right) &= \left(g\left(x\right)/K^{\frac{1}{\alpha}}\right)^\alpha \left(\phi\left(\left(g\left(x\right)\right)^2/K^{\frac{2}{\alpha}}\right)\right)^{\frac{1}{2}} \\ &\le \left(g\left(x\right)\right)^\alpha \left(\phi\left(g\left(x\right)\right)^2\right)^{\frac{1}{2}}/K \\ &= x/K \text{ (by (310))} \\ &= f\left(g\left(x/K\right)\right). \end{aligned} \tag{311}$$

Hence

$$g\left(x\right)/K^{\frac{1}{\alpha}} \le g\left(x/K\right). \tag{312}$$

Let $D = 1/K^{2/\alpha}$. By (312) and (309) , we obtain

$$\mathrm{E}\exp\left(D\left(g\left(|X|\right)\right)^{2}\right) \le \mathrm{E}\exp\left(\left(g\left(|X|/K\right)\right)^{2}\right) \le M, \tag{313}$$

as desired.

The following is a link between the combinatorial structure of $F \subset \mathbf{R}^n$ and tail probability estimates involving random variables in $L_F^2\left(\Omega^n, P^n\right)$..

Theorem (5.2.22)[218]: For $n \in N, F \subset \mathbf{R}^n$, and α -Orlicz function Φ,

$$d_F\left(\Phi\right) < \infty \Leftrightarrow \sup\left\{\|X\|_{\psi_\Phi} : X \in B_{L_F^2\left(\Omega^n\right)}\right\} < \infty \tag{314}$$

Proof: Observe that statement (iv) in Lemma (5.2.19) is the same as statement (i) in Lemma (5.2.19).Then by Lemma (5.2.19) and Lemma (5.2.21)

$$\sup\left\{\|X\|_{L^p} / p^{\frac{\alpha}{2}}\left(\phi(p)\right)^{1/2} : p > 2.X \in B_{L_F^2\Omega^n}\right\} < \infty$$

$$\Leftrightarrow \sup\left\{\|X\|_{\psi_\Phi} : X \in B_{L_F^2\Omega^n}\right\} < \infty. \tag{315}$$

Because $\quad \Phi^{1/2}\left(p\right) = p^{\alpha/2}\left(\phi\left(p\right)\right)^{1/2}$

$$\eta_F\left(\Phi^{1/2}\right) = \sup\left\{\|X\|_{L^p} / p^{\alpha/2}\left(\phi\left(p\right)\right)^{1/2} : p > 2, X \in B_{L_F^2\left(\Omega^n\right)}\right\} \tag{316}$$

(Definition(5.2.16)) Hence

$$\eta_F\left(\Phi^{\frac{1}{2}}\right) < \infty \Leftrightarrow \sup\left\{\|X\|_{\psi_\Phi} : X \in B_{L_F^2\left(\Omega^n\right)}\right\} < \infty, \tag{257}$$

which, by Corollary (5.2.18) implies (314).

Corollary (5.2.23)[274]: If $n \in N$ *and* $F \subset \mathbf{R}^n$,then for α -Orlicz functions $\mathbf{\Phi}_j$,

$$d_F\left(\sum_{j=1}^{n}\mathbf{\Phi}_j\right) < \infty \text{ if and only if } \sup\left\{\sum_{j=1}^{n}\|X_j\|_{\psi_{\Phi_j}} : X \in B_{L_F^2\left(\Omega^n\right)}\right\} < \infty \tag{318}$$

Proof: By Lemma (5.2.19) and Lemma (5.2.21) we have

$$\sup\left\{\sum_{j=1}^{n}\left\|X_j\right\|_{L^p} / p^{\frac{\alpha}{2}}\left(\phi_j\left(p\right)\right)^{1/2} : p > 2, X_j \in B_{L_F^2\Omega^n}\right\} < \infty$$

if and only if $\sup\left\{\sum_{j=1}^{n}\left\|X_j\right\|_{\psi_{\Phi_j}} : \text{for } X_j \in B_{L_F^2\Omega^n}\right\} < \infty.$ (319)

since $\sum_{j=1}^{n}\Phi_j^{1/2}\left(p\right) = \sum_{j=1}^{n} p^{\alpha/2}\left(\phi_j\left(p\right)\right)^{1/2}$, *then*

$$\eta_F\left(\sum_{j=1}^{n}\Phi_j^{1/2}\right) = \sup\left\{\sum_{j=1}^{n}\left\|X_j\right\|_{L^p} / p^{\alpha/2}\left(\sum_{j=1}^{n}\phi_j\left(p\right)\right)^{1/2} : p > 2, X_j \in B_{L_F^2\left(\Omega^n\right)}\right\}.$$

Hence

$$\eta_F\left(\sum_{j=1}^{n}\Phi_j^{1/2}\right) < \infty \text{ if and only if } \sup\left\{\sum_{j=1}^{n}\left\|X_j\right\|_{\psi_{\Phi_j}} : \text{for } X_j \in B_{L_F^2\Omega^n}\right\} < \infty$$

Which gives the result by Corollary (5.2.18).

Chapter 6

Mass Transportation of Free Functional Inequalities and Poincare Inequalities

We permit to give a new and very short proof of a result of Otto and Villani-Generalization to other type of concentration are also considered. In particular, we show that the Poincar'e inequality is equivalent to a certain form of dimension free exponential concentration. The proofs of these result rely on simple large Deviations techniques. We give equivalent functional form of these Poincare type inequalities in terms of transportation-cost inequalities and inf-convolution inequalities workable sufficient conditions are given a comparison is made with super Poincar'e inequalities , we also addresses two version of free Poincare inequalities and their interpretation in terms of spectral properties of Jacobi operators. The last establish the corresponding inequalities for measures on $\mathbf{R}_+$ with the reference example of the Marcenko-pastar distribution. We show some verifications of series of transportations inequalities. We give a result by using a nondecreasing super additive function. Wegeneralize a Lamma used in deriving concentration inequalities and Bobkov-Ledoux result. We determined a particular value of delta with a general some potential. We find a norm of a projection with respect to $-1 \le \lambda \le 1$. We deduce the values of $W\left(\mu, \mu_V\right)$

interms of the relative free Fisher Information to construct the transportation cost result.

Section (6.1): Characterization of Dimension Free Concentration In Transportation Inequality

One says that a probability measure on $\mathbf{R}^d$ has the Gaussian dimension free concentration property if there are three non-negative constants a, b and r_0 such that for every integer n, the product measure μ^n verifies the following inequality:

$$\mu^n\left(A+rB_2\right)\geq 1-be^{-a(r-r_0)^2} \quad \forall r\geq r_0 \tag{1}$$

For all measurable subset A of $\left(\mathbf{R}^d\right)^n$ with $\mu^n(A)\geq\frac{1}{2}$ denoting by B_2 the Euclidean unit ball of $\left(\mathbf{R}^d\right)^n$.The first example is of course the standard Gaussian measure on $\mathbf{R}$ for which the inequality (1) holds true with the sharp constants $r_0=0, a=\frac{1}{2}$ and $b=\frac{1}{2}$. Gaussian concentration is no the only possible behavior; for example, if $\rho\in[1,2]$ the probability measure $d\mu_p(x)=Z_p^{-1}e^{-|x|^p}dx$ verifies a concentration inequality similar to (1) with r^2 replaced by $\min\left(r^p,r^2\right)$. In recent years many developed various functional approaches to the concentration of measure phenomenon. For example, the logarithmic-Sobolev inequality is well known to imply (1); this is renowned Herbst argument[174],[165,255],[276,254],[252,101,61],[138,179,217.257,69,256, 191],[17],[238],[261,239,58,59],[120]and[23,22] .

One shows with a certain generality that Talagrand's transportation-cost inequalities are equivalent to dimension free concentration of measure. Let us give a flavor of our results in the Gaussian case. Let us first define the optional quadratic transportation-cost on $P\left(\mathbf{R}^d\right)$, one defines

$$T_2(\mathrm{V},\mu)=\inf_{\pi}\int|x-y|_2^2\,d\pi(x,y) \tag{2}$$

where π describes the set $P(v,\mu)$ of probability measures on $\mathbf{R}^d\times\mathbf{R}^d$ having V and μ for marginal distributions. One says that μ verifies the inequality $T_2(C)$, if

$$T_2(\mathrm{V},\mu)\leq CH(\mathrm{V}|\mu),\quad \forall\mathrm{V}\in\mathrm{P}\left(\mathbf{R}^d\right) \tag{3}$$

Where $H(\mathrm{V}|\mu)$ is the relative entropy of V with respect to μ defines by

$H(\mathrm{V}|\mu)=\int \log\left(\frac{d\mathrm{V}}{d\mu}\right)d\mathrm{V}$ if V is absolutely continuous with respect to μ and $+\infty$ otherwise. The idea of controlling an optimal transportation-cost inequalities by the relative entropy to obtain concentration first appeared in Marton's works [138,139]. The inequality T_2 was then introduced by Talagrand in [179], where it was proved to be fulfilled b Gaussian probability measures in particular, if $\mu=\gamma$ is the standard Gaussian measure on $\mathbf{R}$, then the inequality (3) holds true with the sharp constant $\mathrm{C}=2$. We show theorem .

Theorem (6.1.1) [194]: Let μ be a probability measure on $\mathbf{R}^d$ and $a>0$; the following propositions are equivalent:

(i) There are $r_0, b\geq 0$ such that for all n the probability μ^n verifies (1).

(ii) The probability measure μ verifies $T_2(1/a)$.

Let μ be a probability measure on χ and $(X_i)^i$ an $i.i.d$ sequence of random variables with law μ defined on some probability space (Ω,P). The empirical measure L_n is defined for all integer n by $L_n=\frac{1}{n}\sum_{i=1}^{n}\delta_{X_i}$, where δ_x, stands for the Dirac mass at point x.

According to Varadarajan's Theorem (see [222]), with probability 1 the sequence $(L_n)_n$ converges to μ in $\mathrm{P}(\chi)$ for the topology of weak convergence, this means that there is a measurable subset N of Ω with $\mathrm{P}(N)=0$ such that for all $w\notin N$,

$$\int f dL_n(w)\xrightarrow[n\to+\infty]{}\int f d\mu$$

for all bounded continuous f on X.

The topology of weak convergence can be metrized by various metrics. Here, one will consider the Wasserstein metrics. Let $p\geq 1$ and define

$$\mathrm{P}_p(\chi)=\left\{\mathrm{V}\in\mathrm{P}(\chi):\int \rho(x_0,x_1)^p\, d\mathrm{V}(x)<\infty,\right\}\text{ for some } x_0\in\chi .$$

For all probability measures $\mathrm{V}_1,\mathrm{V}_2\in\mathrm{P}_p(\chi)$, define

$$T_p(\mathrm{V}_1,\mathrm{V}_2)=\inf_{\pi}\int \rho(x,y)^p d\pi(x,y) \text{ and } \mathrm{W}_p(\mathrm{V}_1,\mathrm{V}_2)=\left(T_p(\mathrm{V}_1,\mathrm{V}_2)\right)^{\frac{1}{p}}$$

where π describes the set $\mathrm{P}(\mathrm{V}_1,\mathrm{V}_2)$ of couplings of V_1 and V_2.

According to [23] W_p is a metric on $P_p(\chi)$ and for every sequence μ_n in $P_p(\chi)$, $W_p(\mu_n,\mu)\to 0$, if and only if μ_n converges to μ, for the weak topology and $\int \rho(x_0,x)^p\, d\mu_n \to \int \rho(x_0,x)^p\, d\mu$, for some (and thus any) $x_0 \in \chi$.

From these considerations, one can conclude that if $\mu \in P_p(\chi)$ then $W_p(L_n,\mu)\to 0$ with probability one, and in particular, $P\left(W_p(L_n,\mu)\geq t\right)\to 0$ when $n\to+\infty$, for all $t>0$. Moreover, supposing that $\mu\in P_p(\chi)$ with $p>1$, it is easy to check that the sequence $W_p(L_n,\mu)$ is bounded in $L_p(\Omega,P)$, thus it is uniformly integrable and consequently $E\left[W_p(L_n,\mu)\right]\to 0$. This is summarized in the following proposition:

Proposition (6.1.2) [194]: If $\mu\in P_p(\chi)$, then the sequence $W_p(L_n,\mu)\to 0$ almost surely (and thus in probability) and if $p>1$, then the convergence is in $L_1 : E\left[W_p(L_n,\mu)\right]\to 0$.

On the other hand, Sanov's Theorem (see [5]) says that for all good sets $A, P(L_n\in A)$, behaves like $e^{-nH(A|\mu)}$ when n is large, where $H(A|\mu)$ stands for the infimum of $H(.|\mu)$ on A. So, when A does not contain $\mu, H(A|\mu)>0$ and this probability tends to 0 exponentially fast. With this in mind, one can expect that $P\left(W_p(L_n,\mu)>t\right)$ behaves like $e^{-nH(t)}$ where $H(t)=\inf\left\{H(V|\mu): V\ s.t.\ W_p(V,\mu)>t\right\}$.

The following result validates partially this heuristic, stating that $P\left(W_p(L_n,\mu)>t\right)$ tends to 0 not faster than $e^{-nH(t)}$.

As in [190], the use of this Large Deviations technique will be the key step in the proof of Theorem (6.1.1) .

As in the preceding section, (χ,ρ) will be a Polish space. The product space χ^n will be equipped with the following metric:

$$\rho_2^n(x,y)=\left[\sum_{i=1}^{n}\rho\left(x^i,y^i\right)^2\right]^{\frac{1}{2}}$$

(here $x=\left(x^1,x^2,\ldots,x^n\right)$ with $x^i\in\chi$ for all i).

In the general case, one says that a probability measure μ on (χ,ρ) verifies the dimension free Gaussian concentration property, if there are $r_0,a,b\geq 0$

such that for all n the probability μ^n verifies

$$\mu^n\left(A^r\right) \geq 1 - be^{-a(r-r_0)^2}, \quad \forall r \geq r_0 \tag{4}$$

for all measurable $A \subset \chi^n$ such that $\mu_n(A) \geq \frac{1}{2}$, where A^r denotes the r-enlargement of A defined by $A^r = \{ x \in \chi^n$ such that there is $\bar{x} \in A$ with $\rho_2^n(x, \bar{x}) \leq r \}$

Of course, when $\chi = \mathbf{R}^d$ is equipped with its Euclidean metric one has $A^r = A + rB_2$ and one recovers the inequality (1).

Let us recall the inequality of the T_1 transportation-cost inequality. One says that a probability measure μ on χ verifies $T_1(C)$, if

$$W_1(V, \mu) \leq \sqrt{CH(V|\mu)}, \quad \forall V \in P(\chi)$$

According to Jensen's inequality, the inequality $T_1(C)$ is weaker than $T_2(C)$; it was completely characterized in terms of square exponential integrability in [93].

The proof of the following well known result makes use of the so called Marton's argument.

Proposition (6.1.3) [194]: (Marton). If μ verifies $T_1(C)$, then for all measurable subset A of χ, such that $\mu(A) \geq \frac{1}{2}$.

$$\mu\left(A^r\right) \geq 1 - be^{-c^{-1}(r-r_0)^2}, \quad \forall r \geq r_0$$

where $r_0 = \sqrt{C \log(2)}$.

Proof: Consider a subset A of χ, and define $d\mu_A = 1_A d\mu(x)/\mu A$. Let $B = \chi \setminus A^r$ and define μ_B accordingly. Since the distance between two points of A and B is always more than r, one has $W_1(\mu_A, \mu_B) \geq r$. The triangle inequality and the transportation-cost inequality $T_1(C)$ yield[274]

$$r \le W_1(\mu_A,\mu_B) \le W_1(\mu_A,\mu) + W_1(\mu_B,\mu) \le \sqrt{CH(\mu_A|\mu)} + \sqrt{CH(\mu_B|\mu)}$$
$$= \sqrt{C\log(1/\mu(A))} + \sqrt{C\log(1/\mu(B))} = \sqrt{c\log 1 - c\log\mu(A)} + \sqrt{c\log 1 - c\log\mu(B)}$$
$$= \sqrt{-c\log\mu(A)} + \sqrt{-c\log\mu(B)} \le \sqrt{-c\log\frac{1}{2}} + \sqrt{-c\left[\log\mu(X) - \log\mu(A^r)\right]}$$

$$= \sqrt{-c\log 1 + c\log 2} + \sqrt{-c\log\mu(x) + c\log\mu(A^r)} = \sqrt{c\log 2} + \sqrt{-c\left[\log\mu(X) + \log\mu(A^r)\right]}$$

$$= r_0 + \sqrt{-c\left[\log 1 + \log\mu(A^r)\right]} \Rightarrow r - r_0 \le \sqrt{-c\left[\log 1 - \log\mu(A^r)\right]}$$

$$\Rightarrow (r-r_0)^2 \le -c\left[\log 1 - \log\mu(A^r)\right] \Rightarrow -\frac{1}{c}(r-r_0)^2 \le \left(\log 1 - \log\mu(A^r)\right)$$

$$\left[e^{-\frac{1}{c}}(r-r_0)^2 \ge 1 - \mu(A^r) \Rightarrow \mu(A^r) \ge 1 - e^{-1/2(r-r_0)^2}\right.$$

Rearranging terms gives the result.

Theorem (6.1.4) [194]: Let $\mu \in P_2(\chi)$ and $a > 0$; the following Propositions are quivalent:

(i) There are $r_0, b \ge 0$ such that for all n the probability μ^n verifies (4),

(ii) The probability μ verifies $T_2(1/a)$.

Proof: Let us show that (ii) implies (i). The main point is that T_2 tensorizes ; this means that if μ verifies $T_2(1/a)$ then μ^n verifies $T_2(1/a)$ on the space χ^n equipped with ρ_2^n we can find a general result concerning tensorization properties of transportation-cost inequalities in [190]. Jensen's inequality implies that $W_1^2 \le T_2$ and consequently μ^n verifies $T_1(1/a)$(on χ^n equipped with ρ_2^n) for all n. Applying Proposition (6.1.3) to μ^n gives (1) with $r_0 = \sqrt{\log(2)/a}$, $b = 1$ and a.

Let us show that (i) implies (ii). For every integer n, and $x \in \chi^n$, define $L_n^x = n^{-1}\sum_{i=1}^n \delta_{x^i}$. The map $x \mapsto W_2(L_n^x, \mu)$ is $1/\sqrt{n} - Lipschitz$. Indeed, if $x = (x^1,\ldots,x^n)$ and $y = (y^1,\ldots,y^n)$ are in χ^n, then thanks to the triangle inequality,

$$\left|W_2\left(L_n^x,\mu\right)-W_2\left(L_n^y,\mu\right)\right|\le W_2\left(L_n^x,L_n^y\right)$$

According to the convexity property of $T_2(.,.)$ (see e.g. [257]), we have

$$T_2\left(L_n^x,L_n^y\right)\le\frac{1}{n}\sum_{i=1}^{n}T_2\left(\delta_{x^i},\delta_{y^i}\right)=\frac{1}{n}\sum_{i=1}^{n}\rho\left(x^i,y^i\right)^2=\frac{1}{n}\rho_2^n\left(x,y\right)^2$$

which proves the claim.

Now, let $(X_i)_i$ be an i.i.d sequence of law μ and let L_n be its empirical measure. Let m_n be the median of $W_2(L_n,\mu)$ and define $A=\left\{x:W_2\left(L_n^x,\mu\right)\le m_n\right\}$. Then $\mu^n(A)\ge 1/2$ and it is easy to show that $A^r\subset\left\{x:W_2\left(L_n^x,\mu\right)\le m_n+r/\sqrt{n}\right\}$. Applying (4) to A gives

$$P\left(W_2\left(L_n,\mu\right)>m_n+r\sqrt{n}\right)\le b\exp\left(-a\left(r-r_0\right)^2\right),\quad\forall r\ge r_0$$

Equivalently, as soon as $\sqrt{n}\left(u-m_n\right)\ge r_0$, one has

$$P\left(W_2\left(L_n,\mu\right)>u\right)\le b\exp\left(-a\left(\sqrt{n}\left(u-m_n\right)-r_0\right)^2\right).$$

Now, since $W_2(L_n,\mu)$ converges to 0 in probability (see Proposition (6.1.2)), the sequence m_n goes to 0 when n goes to $+\infty$. Consequently,

$$\log P\left(W_2\left(V,\mu\right)>\mu\right)=\lim_{n\to\infty}\frac{1}{n}\log P\left(W_1\left(L_n,\mu\right)>\mu\right)\le$$

$$\limsup_{n\to+\infty}\frac{1}{n}\log P\left(W_2\left(L_n,\mu\right)>u\right)\le -au^2,\quad\forall u>0$$

The final step is given by Large Deviations. According to Theorem (6.1.4),

$$\limsup_{n\to+\infty}\log P\left(W_2\left(L_n,\mu\right)>u\right)\ge-\inf\left\{H\left(V|\mu\right):V\in P_2\left(\chi\right)s.t.W_2\left(V,\mu\right)>u\right\}$$

This together with the preceding inequality yields

$$\inf\left\{H\left(V|\mu\right):V\in P_2\left(\chi\right)s.t.W_2\left(V,\mu\right)>u\right\}\ge au^2$$

or in other words,

$$aW_2\left(V,\mu\right)^2\le H\left(V|\mu\right)$$

and this achieves the proof.

Let us make a remark on the proof. We will notice that the second part of the proof applies if one replaces $W_2(.,\mu)$ by any application $\Phi : P(\chi) \to \mathbf{R}^+$ which is continuous with respect to the weak top'ology, verifies $\Phi(\mu)=0$, and is such that for all integer n, the map $\chi^n \to \mathbf{R}^+ : x \mapsto \Phi(L_n^x)$ is $1/\sqrt{n} - Lipschitz$ for the metric ρ_2^n on χ^n. For such an application Φ, one can show, with exactly the same proof, that the dimension free Gaussian concentration property (4) implies that $a\Phi^2(V) \le H(V|\mu)$, for all V and it could be that this new inequality is stronger than T_2. Actually; it is not the case. Namely, it is to show that if Φ verifies the above listed properties ,then $\Phi(V) \le W_2(V,\mu)$, for all V, and so the choice $\Phi = W_2$ is optimal.

Our aim is now to recover and extend a theorem by Otto and Villani stating that the Logarithmic-Sobolev inequality is stronger than Talagrand's T_2 inequality.

Let us recall that a probability measure μ on χ verifies the Logarithmic-Sobolev inequality with constant $C>0$ (LSI(C) for short) if

$$H_\mu\left(f^2\right) \le C\int |\nabla|^2 d\mu$$

for all locally Lipschitz f, where the entropy functional is defined by

$$H_\mu(f) = \int f \log f d\mu - \int f d\mu \log\left(\int f d\mu\right), \quad f \ge 0,$$

and the length of the gradient is defined by

$$|\nabla f|(x) = \limsup_{y\to x} \frac{|f(x)-f(y)|}{\rho(x,y)} \tag{5}$$

(when X is an isolated point, we put $|\nabla f|(x)=0$.

. Namely, if μ, verifies the $LSI(C)$ inequality, then according to the additive property of the Logarithmic-Sobolev inequality, one can conclude that the product measure μ^n verifies

$$H_{\mu^n}\left(f^2\right) \le C\int \sum_{i=1}^{n} |\nabla_i f|^2(x)\, d\mu^n(x) \tag{6}$$

where the length of the 'partial derivative' $|\nabla_i f|$ is defined according to (5).

The problem is that, in this very abstract setting, $\sum_{i=1}^{n} |\nabla_i f|^2(x)$ and $|\nabla f|^2(x)$ (computed with respect to ρ_2^n) may be different. The tensorized Logarithmic-Sobolev inequality will yield concentration inequalities for functions such that $\sum_i |\nabla_i f|^2(x) \le 1 \mu^n -$almost everywhere and this class of functions may not contain 1-Lipschitz functions for the ρ_2^n metric. Nevertheless, this difficulty can be circumvented as shown in the following theorems.

Theorem (6.1.5) [194]: Let μ, be a probability measure on χ and suppose that for all integer n the function F_n defined on χ^n by $F_n(x) = \mathrm{W}_2(L_n^x, \mu)$ verifies

$$\sum_{i=1}^{n} |\nabla_i F_n|^2(x) \le 1/n, \textit{ for } \mu^n \text{ almost every } x \in \chi^n. \tag{7}$$

If μ, verifies the inequality $LSI(C)$, then μ, verifies the inequality $T_2(C)$. We have seen during the proof of Theorem (6.1.4) that the functions F_n are $1/\sqrt{n} - Lipschitz$ for the metric ρ_2^n. Suppose that $\chi = \mathbf{R}^d$ or a Riemannian manifold M, then according to Rademacher's Theorem, F_n is almost everywhere differentiable on $(\mathbf{R}^d)^n$ (resp. M^n) with respect to the Lebesgue measure. It is thus easy to show that condition (188) is fulfilled when μ, is absolutely continuous with respect to Lebesgue measure. This permits us to recover Otto and Villani's result as stated in [69].

Proof: As we said above the product measure μ^n verifies the inequality (6). Apply this inequality to $f = e^{\frac{s}{2}F_n}$, with $s \in \mathbf{R}^+$. It is easy to show that $\left|\nabla_i e^{\frac{s}{2}F}\right| = \frac{s}{2} e^{\frac{s}{2}F_n} |\nabla_i F_n|$, thus, using condition (7), one sees that the right hand side of (6) is less than $C\frac{s^2}{4n}\int e^{sF_n} d\mu^n$.

Letting $\mathrm{Z}(s) = \int e^{sF_n} d\mu^n$, one gets the differential inequality:

$$H\mu^n(f^2) \le C\int \sum_{i=1}^{n} |\nabla_i f^2|(x) d\mu^n(x) = C\int \sum_{i=1}^{n} \left|\nabla_i e^{\frac{s}{2}Fn}\right|^2 (x) d\mu^n(x)$$

$$=C\int\sum_{i=1}^{n}\frac{s^2}{4}e^{sFn}\left|\nabla_i Fn\right|^2(x)d\mu^n(x)\le\frac{Cs^2}{4n^2}\int e^{sFn}d\iota\mu^n(x)$$

Since $\left|\nabla_i e^{\frac{s}{2}Fn}\right|=\frac{s}{2}e^{\frac{s}{2}Fn}\left|\nabla_i Fn\right|$

From the definition of the relative entropy we deduce that

$$H\left(\mu^n\middle|f^2\right)=\int f^2\log f^2 d\mu^n-\int f^2 d\mu^n\log\left(\int f^2 d\mu^n\right)$$

$$=\int sFne^{sFn}d\beta\mu^n-\int e^{sFn}d\mu^n\log\left(\int e^{sFn}d\mu^n\right)$$

Then we see that $\int sFne^{sFn}d\mu^n-\int e^{sFn}d\mu^n\log\left(\int e^{sFn}d\mu^n\right)\le\frac{Cs^2}{4n}\int e^{sFn}d\mu^n$

Letting $Z(s)=\int e^{sFn}d\mu^n$ We get the differential inequality

$sz'(s)-z(s)\log z(s)\le\frac{Cs^2}{4n}z(s)$.So that $\frac{Z'(s)}{sZ(s)}-\frac{\log Z(s)}{s^2}\le\frac{\mathrm{C}}{4n}$

The integrating this yields

$$z(s)=\int e^{sFn}d\mu^n\le e^s\int Fnd\mu^n+\frac{Cs^2}{4},\qquad\forall s\in\mathbf{R}^+$$

This implies that

$$\mathrm{P}\left(\mathrm{W}_2(L_n,\mu)\ge t+\mathrm{E}\left[\mathrm{W}_2(L_n,\mu)\right]\right)\le e^{-nt\,2/\mathrm{C}}$$

According to Proposition (6.1.2)$\mathrm{E}\left[\mathrm{W}_2(L_n,\mu)\right]\to 0$. Arguing exactly as in proof of Theorem (6.1.4), one concludes that the inequality $T_2(C)$ holds.

With an extra assumption on the support of μ, one shows in the following theorem that the implication $LSI\Rightarrow T_2$ is true with a relaxed constant:

Theorem (6.1.6) [194]: Let μ be a probability measure on χ such that

$$\mu\left\{x\in\chi\ s.t.\rho^2(x,u)-\rho^2(x,\mathrm{V})=K\right\}=0,\forall K\in\mathbf{R},\forall u\ne\mathrm{V}\in\chi\qquad(8)$$

If μ verifies the inequality $LSI(C)$ then μ satisfies $T(2C)$.

The condition (8) first appeared in a paper by Cuesta-Albertos and Tuero-

Dfaz on optimal transportation. Roughly speaking, this assumption guaranties the uniqueness of the Monge-Kantorovich Problem of transporting μ on a probability measure V with finite support (see [106]). For μ on $\mathbf{R}^d$, the condition (8) amounts to say that μ does not charge hyperplanes. We think that working better it would be possible to obtain the right constant C instead of 2C.

Proof: We will use a sort of symmetrization argument. First observe that the probability measure $\mu^n \times \mu^n$ verifies the following Logarithmic-Sobolev inequality:

$$H\left(\mu^n \times \mu^n \left| f^2\right.\right) \le C\sum_{i=1}^{n} \left|\nabla_{i,1} f\right|^2 (x,y) + \left|\nabla_{i,2} f\right|^2 (x,y)\, d\mu^n(x)\, d\mu^n(y)$$

for all $f : \chi^n \times \chi^n \to \mathbf{R} : (x,y) \mapsto f(x,y)$, where $\left|\nabla_{i,1} f\right|$(resp. $\left|\nabla_{i,2} f\right|$) denotes the length of the gradient with respect to the x^i-coordinate (resp. the y^i-coordinate).

Define $G_n(x,y) = \mathrm{W}_2\left(L_n^x, L_n^y\right)$ for all $x, y \in \chi^n$. One wants to apply the tensorized Logarithmic-Sobolev inequality to the function G_n. To do so one needs to compute the length of its partial derivatives. Let us explain how to compute $L = \left|\nabla_{1,1} G_n\right|(a,b)$, for instance. For every $z \in \chi$, let $za = \left(z, a^2, ..., a^n\right)$; obviously,

$$L = \limsup_{z\to a^1} \frac{\left|\mathrm{W}_2\left(L_n^{za}, L_n^b\right) - \mathrm{W}_2\left(L_n^a, L_n^b\right)\right|}{\rho\left(z, a^1\right)} = \frac{1}{2\mathrm{W}_2\left(L_n^a, L_n^b\right)} \limsup_{z\to a^1} \frac{\left|T_2\left(L_n^{za}, L_n^b\right) - T_2\left(L_n^a, L_n^b\right)\right|}{\rho\left(z, a^1\right)}$$

According to the condition (8), the probability measure μ is diffuse; so the probability of points $x \in \chi^n$ having distinct coordinates is one. So, one can suppose without restriction that the coordinates of a (resp. b) are all different. If z is sufficiently close to a^1, the coordinates of z and a are all distinct too. According to e.g [23], the optimal transport of L_n^a on L_n^b is given by a permutation, this means that there is at least one permutation a of $\{1,...,n\}$ such that

$$T_2\left(L_n^a, L_n^b\right) = n^{-1} \sum_{i=1}^{n} \rho\left(a^i, b^{\sigma(i)}\right)^2$$

Let us denote by S the set of these permutations and define accordingly the set S_z of permutations realizing the optimal transport of L_n^{za} on L_n^b.

Without loss of generality, one can suppose that S is a singleton. Indeed, let σ and $\tilde{\sigma}$ be two distinct permutations and consider

$$H_{\sigma,\tilde{\sigma}} = \left\{ x \in \chi^n : \sum_{i=1}^{n} \rho\left(x^i, b^{\sigma(i)}\right)^2 = \sum_{i=1}^{n} \rho\left(x^i, b^{\sigma(i)}\right)^2 \right\}$$

Applying Fubini's Theorem together with the condition (8), one gets easily that $\mu^n\left(H_{\sigma,\tilde{\sigma}}\right) = 0$. This readily proves the claim. In the sequel we will set $S = \left\{\sigma^*\right\}$.

Now we claim that if z is sufficiently close to a^1, then $S_z = \left\{\sigma^*\right\}$. Indeed, let

$$\varepsilon_0 = \min_{\sigma \neq \sigma} \left\{ n^{-1} \sum_{i=1}^{n} \rho\left(a^i, b^{\sigma(i)}\right)^2 - T_2\left(L_n^a, L_n^b\right) \right\} > 0$$

then there is a neighborhood V of a^1 such that for all $z \in V$, one has

$$\left| T_2\left(L_n^{za}, L_n^b\right) - T_2\left(L_n^a, L_n^b\right) \right| \leq \varepsilon_0 / 3$$

and for all permutation σ,

$$\left| n^{-1} \sum_{i=1}^{n} \rho\left((za)^i, b^{\sigma(i)}\right)^2 n^{-1} \sum_{i=1}^{n} \rho\left(a^i, b^{\tilde{\sigma}(i)}\right)^2 \right| \leq \varepsilon_0 / 3$$

Now, if $z \in V$ and $\sigma \in S_z$, one has

$$\left| n^{-1} \sum_{i=1}^{n} \rho\left(a^i, b^{\sigma(i)}\right)^2 \leq n^{-1} \sum_{i=1}^{n} \rho\left((za)^i, b^{\sigma(i)}\right)^2 \right| + \varepsilon_0 / 3$$

$$= T_2\left(L_n^{za}, L_n^b\right) + \varepsilon_0 / 3 \leq \mathrm{T}_2\left(L_n^a, L_n^b\right) + 2\varepsilon_0 / 3$$

By the definition of the number co, one concludes that $\sigma = \sigma^*$, which proves the claim. Now, if $z \in V$, then

$$\frac{\left| T_2\left(L_n^{za}, L_n^b\right) - T_2\left(L_n^a, L_n^b\right) \right|}{\rho\left(z, a^1\right)} = \frac{\left| \rho\left(z, b^{\sigma^{*(1)}}\right)^2 - \rho\left(a^1, b^{\sigma^{*(1)}}\right)^2 \right|}{n\rho\left(z, a^1\right)}$$

$$= \frac{1}{np\left(z, a'\right)} \left| \rho\left(z, b^{\sigma(1)}\right) - \rho \left| a, b^{\sigma(1)} \right| \rho\left(z, b^{0^{*(1)}}\right) + \rho\left(a, b^{\sigma(1)}\right) \right|$$

$$\leq \frac{\rho\left(z, a'\right)}{n\rho\left(z, a^1\right)} \left| \rho\left(z, b^{\sigma^*(1)}\right) + \rho\left(a, b^{\sigma^*(1)}\right) \right| = \frac{1}{n} \left(\rho\left(z, b^{\sigma^*(1)}\right) + \rho\left(s, b^{\sigma^*(1)}\right) \right)$$

$$\leq \frac{1}{n}\rho\left(z, b^{\sigma^*(1)}\right)+\rho\left(a^1, b^{\sigma^*(1)}\right)$$

So letting $z \to a^1$, Yields $L \leq \frac{\rho\left(a^1, b^{\sigma^*(1)}\right)}{n\mathrm{W}_2\left(L_n^a, L_n^b\right)}$.

Doing the same for the other partial derivatives yields:

$$\sum_{i=1}^{n}\left|\nabla_{i,1}G_n\right|^2(a,b) \leq \frac{\sum_{i=1}^{n}\rho\left(a^i, b^{\sigma^*(i)}\right)^2}{n^2 T_2\left(L_n^a, L_n^b\right)} = \frac{1}{n}$$

Finally,

$$\sum_{i=1}^{n}\left|\nabla_{i,1}G_n\right|^2(a,b)+\left|\nabla_{i,2}G_n\right|^2(a,b) \leq \frac{2}{n}$$

for $\mu^n \times \mu^n$ almost every $a, b \in \chi^n \times \chi^n$.

Now reasoning as in the proof of Theorem (6.1.14), one concludes that

$$\mathrm{P}\left(\mathrm{W}_2\left(L_n^X, L_n^Y\right) > t + \mathrm{E}\left[\mathrm{W}_2\left(L_n^X, L_n^Y\right)\right]\right) \leq e^{-nt^{2/(2C)}}$$

On the other hand, an easy adaptation of Theorem (6.1.14) yields

$$\liminf_{n\to+\infty}\frac{1}{n}\log \mathrm{P}\left(\mathrm{W}_2\left(L_n^X, L_n^Y\right) > t + \mathrm{E}\left[\mathrm{W}_2\left(L_n^X, L_n^Y\right)\right]\right)$$
$$-\inf\left\{H\left(\mathrm{V}_1|\mu\right)+H\left(\mathrm{V}_2|\mu\right) : \mathrm{V}_1, \mathrm{V}_2 \in \mathrm{P}_2(\chi)\ s.t. W_2\left(\mathrm{V}_1, \mathrm{V}_2\right) > t\right\}$$

From this follows as before that

$$T_2\left(\mathrm{V}_1|\mu\right) \leq 2C\left(H\left(\mathrm{V}_1|\mu\right)+H\left(\mathrm{V}_2|\mu\right)\right)$$

holds for all probability measures $\mathrm{V}_1, \mathrm{V}_2$ belonging to $\mathrm{P}_2(\chi)$. Taking $\mathrm{V}_2 = \mu$ gives the inequality $T_2(2C)$.

Our next goal is to recover and extend a result of Lott and Villani. Following [117], one says that a probability measure μ on χ verifies the inequality $LSI^+(C)$ if

$$H\left(\mu\left|f^2\right.\right) \leq C\int\left|\nabla^- f\right|^2 d\mu$$

holds true for all locally Lipschitz f, where the subgradient norm $\left|\nabla^- f\right|$ is defined by

$$\left|\nabla^- f\right| = \limsup_{y\to x} \frac{\left[f(y)-f(x)\right]_+}{\rho(x,y)},$$

with $[a]_+ = \max(a,0)$. Since $\left|\nabla^- f\right| \le |\nabla f|$, the inequality $LSI^+(C)$ is stronger than LSI; more precisely, $LSI^+(C) \Rightarrow LSI(C)$.

Theorem (6.1.7) [194]: If μ verifies the inequality $LSI^+(C)$, then μ verifies $T_2(C)$.

This result was first obtained by Lott and Villani using the Hamilton-Jacobi method. This approach forced them to make many assumptions on χ and μ. In particular, in [117] χ was supposed to be a compact length space and a doubling condition was imposed on μ. The result above shows that the implication $LSI^+ \Rightarrow T_2$ is in fact always true. The following proof uses an argument which I learned from Paul-Marie Samson.

Proof: The inequality LSI^+ tensorizes, so μ^n verifies

$$H\left(\mu^n \middle| f^2\right) \le C\int \sum_{i=1}^{n} \left|\nabla_i^- f\right|^2 d\mu^n$$

Take $f = e^{\frac{s}{2}F_n}$, $s \in \mathbf{R}^+$ with $F_n(x) = W_2\left(L_n^x, \mu\right)$. Once again, it is easy to check that $\left|\nabla_i^- e^{\frac{s}{2}F_n}\right| = \frac{s}{2} e^{\frac{s}{2}F_n} \left|\nabla_i^- F_n\right|$ (note that the function $x \mapsto e^{sx}$ is non decreasing). Reasoning as in the proof of Theorem (6.1.5), it is enough to show that $\sum_i \left|\nabla_i^- F_n\right|^2 (x) \le 1/n$ for μ^{n-} almost all $x \in \chi^n$. Let us show how to compute $\left|\nabla_i^- F_n\right|$. Let $z \in X$, $a = \left(a^1, \ldots a^n\right) \in \chi^n$ and set $za = \left(z, a^2, \ldots, a^n\right)$

$$\left|\nabla_i^- F_n\right|(a) = \frac{1}{2F_n(a)} \limsup_{z\to a^1} \frac{\left[T_2\left(L_n^{za}, \mu\right) - T_2\left(L_n^a, \mu\right)\right]_+}{\rho\left(z, a^1\right)}$$

Let $\pi \in P\left(L_n^a, \mu\right)$ be an optimal coupling; it is not difficult to see that one can write $\pi(dx,dy) = \rho(x,dy)\, L_n^a(dx)$, where $\rho\left(a^i, dy\right) = V_i(dy)$ with $V_1, \ldots, V_n$ probability measures on χ such that $n^{-1}\left(V_1 + \ldots + V_n\right) = \mu$. Let $\tilde{\rho}$ be defined as ρ with z in place of a^1; then $\tilde{\pi} = \tilde{\rho}(x,dy)\, L_n^{za}(dy)$ belongs to $P\left(L_n^{za}, \mu\right)$) (but is not necessary optimal). One has

$$T_2\left(L_n^{za},\mu\right)-T_2\left(L_n^{a},\mu\right)\le\int\rho(x,y)^2\,d\tilde{\pi}(x,y)-\int\rho(x,y)^2\,d\pi(x,y)$$
$$=\frac{1}{n}\sum_{i=1}^{n}\int\rho\left((za)^i,y\right)^2 dv_i(y)-\frac{1}{n}\sum_{i=1}^{n}\int\rho\left(a^i,y\right)^2 d\mathrm{V}_i(y)$$
$$=\frac{1}{n}\int\rho(z,y)^2-\rho\left(a^1,y\right)^2 d\mathrm{V}_1(y)\le\frac{1}{n}\rho\left(z,a^1\right)\int\rho(z,y)+\rho\left(a^1,y\right)d\mathrm{V}_1(y)$$

Since the function $x\mapsto[x]_+$ is non decreasing, one has

$$T\frac{\left|T_2\left(L_n^{za},\mu\right)-T_2\left(L_n^{a},\mu\right)\right|_+}{\rho\left(z,a^1\right)}\le\frac{1}{n}\int\rho(z,y)+\rho\left(a^1,y\right)d\mathrm{V}_1(y)$$

Letting $z\to a^1$ yields $\left|\nabla_i^- F_n\right|(a)^2\le\dfrac{\int\rho\left(a^1,y\right)^2 d\mathrm{V}_1(y)}{n^2T_2\left(L_n^a,\mu\right)}$. Doing the same computations for the other derivatives (with the same optional coupling π), we gets

$$\left|\nabla_i^- F_n\right|(a)^2\le\frac{\int\rho\left(a^i,y\right)^2 d\mathrm{V}_i(y)}{n^2T_2\left(L_n^a,\mu\right)}.$$

Summing these inequalities gives $\sum_i\left|\nabla_i^- F_n\right|^2(a)\le 1/n$ for all $a\in\chi^n$, which achieves the proof. The following theorem can be established with exactly the same proof as Theorem (6.1.8).

Theorem (6.1.8) [194]: Let μ be a probability measure on χ, $p\ge 2$ and $a>0$. The following propositions are equivalent:

(i) There are $r_0,b\ge 0$ such that for every n the probability measure μ^n verifies for all A subset of χ^n with $\mu^n(A)\ge\frac{1}{2}$,

$$\mu^n\left(A^r\right)\ge 1-be^{-a\left(r-r_0\right)^p},\quad\forall r\ge r_0\tag{9}$$

where the enlargement A^r is performed with respect to the metric p_ρ^n on χ^n defined by

$$\rho_p^n(x,y)\left[\sum_{i=1}^{n}p\left(x^i,y^i\right)^p\right]^{1/p},\quad\forall x,y\in\chi^n$$

(ii) The probability measure μ verifies the following transportation cost inequality:

$$T_p(\mathrm{V},\mu) \le a^{-1} H(\mathrm{V}|\mu), \quad \forall \mathrm{V} \in \mathrm{P}_\rho(\chi).$$

We want to find the transportation-cost inequality equivalent to Talagrand's two level concentration inequalities which are well adapted to concentration rates between exponential and Gaussian.

Let us say that a probability measure μ on $\mathbf{R}^d$ satisfies a two level dimension free concentration inequality of order $p \in [1,2]$ if there are two non-negative constants a and b such that for every n the inequality

$$\mu^n\left(A + \sqrt{r}B_2 + \sqrt[p]{r}B_p\right) \ge 1 - be^{-ar}, \quad \forall r \ge 0 \tag{10}$$

holds for all measurable subset A of $\left(\mathbf{R}^d\right)^n$ such that $\mu^n(A) \ge \frac{1}{n}$, where B_2 and B_p are the standard unit balls of $\left(\mathbf{R}^d\right)^n$. Inequalities of this form appear in [182], where it is proved that the measure $d\mu_p(x) = \mathrm{Z}_p^{-1} e^{-|x|^p}$, $p \ge 1$ verifies such a bound.

The transportation-cost adapted to this kind of concentration is defined for all probability measures $\mathrm{V}_1, \mathrm{V}_2$ on $\left(\mathbf{R}^{d^n}\right)$ by

$$T_{2,\rho}(\mathrm{V},\mu) = \inf_{\pi \in \mathrm{P}(\mathrm{V}_1,\mathrm{V}_2)} \int \sum_{i=1}^{n} \sum_{j=1}^{d} \alpha_\rho\left(x_j^i - y_j^i\right) d\pi(x,y)$$

Where $\alpha_\rho(u) = \min\left(|u|^2, |u|^\rho\right)$ (here $x = \left(x^1,\ldots,x^n\right)$ with $x^i \in \mathbf{R}^d$ for all i).

The following lemma collects different facts that are needed in the proof**.**

Lemma (6.1.9) [194]: (i) For all $x, y \ge 0, \alpha_p(x+y) \le 2\alpha_p(x) + 2\alpha_p(y)$.

(ii) For all integer $n > 1$ and all probability measures $\mathrm{V}_1, \mathrm{V}_2$ and V_3 on $\left(\mathbf{R}^d\right)^n$,

$$T_{2,p}(\mathrm{V}_1,\mathrm{V}_3) \le 2T_{2,p}(\mathrm{V}_1,\mathrm{V}_2) + 2T_{2,p}(\mathrm{V}_2,\mathrm{V}_3).$$

(iii) For all integer $n > 1$ and all $r \ge 0$, define

$$B_{2,p}(r) = \left\{ x \in \left(\mathbf{R}^d\right)^n : \sum_{i=1}^{n} \sum_{j=1}^{d} \alpha_p\left(x_j^i\right) \le r \right\}.$$

Then for all $p \in [1,2]$,

$$\frac{1}{12}\left(\sqrt{r}B_2 + \sqrt[p]{r}B_p\right) \subset \sqrt{r}B_{2,p}(r) \subset \sqrt{r}B_2 + \sqrt[p]{r}B_p$$

Theorem (6.1.10) [194]: Let μ be a probability measure on $\mathbf{R}^d$ and $p \in [1,2]$. The following propositions are equivalent:

(i) The two level concentration (10) holds for some non-negative a,b independent of n.

(ii) The probability measure μ verifies the transportation-cost inequality

$$T_{2,\rho}(\mathrm{V}, \mu) \le CH(\mathrm{V}|\mu), \quad \forall \mathrm{V} \in \mathrm{P}\left(\mathbf{R}^d\right)$$

for some constant C.

More precisely, if (10) holds for some constants a, b, then the transportation-cost inequality holds with the constant C = 288/ a. Conversely, if the transportation-cost inequality holds for some constant C, then (10) is true for $b = 2$ and $a = 1/(2C)$.

Proof: Let us recall the proof of (ii) implies (i). According to the tensorization property, for all n and all probability measure V on $\left(\mathbf{R}^d\right)^n$,

$$T_{2,p}\left(\mathrm{V}, \mu^n\right) \le CH\left(\mathrm{V}\middle|\mu^n\right)$$

holds. Take A and B in $\left(\mathbf{R}^d\right)^n$ and define $d\mu_A^n = 1_A d\mu / \mu^n(A)$ and $d\mu_B^n = 1_B d\mu / \mu^n(B)$. According to point (ii) of Lemma (6.1.9) and the transportation-cost inequality satisfied by μ^n, we have[274]:

$$\begin{aligned} T_{2,\rho}\left(\mu_A^n, \mu_B^n\right) &\le 2T_{2,\rho}\left(\mu_A^n, \mu^n\right) + 2T_{2,\rho}\left(\mu_B^n, \mu^n\right) \\ &\le 2CH\left(\mu_A^n \middle| \mu^n\right) + 2CH\left(\mu_B^n \middle| \mu^n\right) = -2C\log\left(\mu^n(A)\mu^n(B)\right) \end{aligned}$$

Define

$$c_{2,p}(A,B) = \inf\left\{r \ge 0 \; s.t. \; \left(A + B_{2,p}(r)\right) \cap B \ne \%\right\}$$

then $T_{2,p}\left(\mu_A^n, \mu_B^n\right) \ge c_{2,p}(A,B)$ and so $\mu^n(A)\mu^n(B) \le e^{-c_{2,p}(A,B)/2C}$.

Now, if $\mu^n(A) \ge \frac{1}{2}$ and $B = \left(\mathbf{R}^d\right)^n \setminus \left(A + B_{2,\rho}(r)\right)$,

$$C_{2,\rho}(A,B) \geq -2c\log\left(\mu^n(A)\mu^n(B)\right) = -2c\left[\log\mu^n(A) + \log\mu^n(B)\right]$$

$$\geq -2c\left[\log\frac{1}{2} + \log\mu^n\left(R^d\right)^n\left(A \mid B_{2,p}(1)\right)\right] \geq -2c\left[\log 1 - \log\mu^n\left(A + B_{2,p}(1)\right)\right]$$

$$= -2c\left[-\log 2 + \log\mu^n\left(\mathbf{R}^d\right)^n - \log\mu^n\left(A + B_{2,p}(1)\right)\right]$$

$$= -2c\left[\log\frac{1}{2}.2 - \log\mu^n\left(A + B_{2,p}(1)\right)\right] = -2c\left[\log\frac{1}{2} - \log\frac{\mu^n}{2}\left(A + B_{2,p}(1)\right)\right]$$

$$\Rightarrow -\frac{1}{2c}c_{2,p}(A,B) \geq \log\frac{1}{2} - \log\frac{\mu^n}{2}\left(A + B_{2,p}(1)\right)$$

Taking the logarithms in both sides

$$\Rightarrow e^{-\frac{1}{2c}c_{2,\rho}(A,\mathrm{P})} \geq \frac{1}{2} - \frac{\mu^n}{2}\left(A + B_{2,p}(1)\right) \Rightarrow \mu^n\left(A + B_{2,p}(1)\right) \geq 1 - 2e^{-\frac{1}{2c}c_{2,\rho}(A,B)}$$

Where $\log 2 \geq 0, b = 2, a = \frac{1}{2}$

we have $c_{2,p}(A,B) = r$ and so $\mu^n\left(A + B_{2,p}(r)\right) \geq 1 - 2e^{-r/2c}$. Using point (iii) of Lemma (6.1.9), gives $\mu^n\left(A + \sqrt{r}B_2 + \sqrt[p]{r}B_p\right) \geq 1 - 2e^{-r/2c}$.

We give that the probability measure μ^n on $\left(R^d\right)^n$ satisfies two level dimension free concentration inequality of order $p \in [1,2]$ if there are two non-negative constants $a = \frac{1}{2c}$ and $b = 2$.

Now let us prove the converse. Let $(X_i)_i$ be an i.i.d sequence of law μ and let L_n be its empirical measure. Consider

$A = \left\{x \in \left(\mathbf{R}^d\right)^n \ s.t. T_{2,p}\left(L_n^x, \mu\right) \leq m_n\right\}$ where m_n denotes the median of $T_{2,p}\left(L_n^x, \mu\right)$. According to point (iii) of Lemma (6.1.9) $A + \sqrt{r}B_2 + \sqrt[p]{r}B_p \subset A + 12B_{2,p}(r)$. Let $x \in A + 12B_{2,p}(r)$; there is some $\bar{x} \in A + 12B_{2,p}(r)$ there is some $\bar{x} \in A$ such that

$$\sum_{i=1}^{n}\sum_{j=1}^{d}\alpha_p\left(\frac{x_j^i - \bar{x}_j^i}{12}\right) \leq r$$

(here $x=\left(x^1,x^2,...x^n\right)$ with $x^i \in \mathbf{R}^d$). Since $\alpha_p\left(x/12\right) \geq \alpha_p\left(x\right)/144$, one gets $T_{2,\rho}\left(L_n^x, L_n^{\bar{x}}\right) \leq 144r/n$. According to point (ii) of Lemma(6.1.9)$T_{2,p}\left(L_n^x, \mu\right)$ $\leq 2T_{2,p}\left(L_n^x, L_n^{\bar{x}}\right) \quad +2T_{2,p}\left(L_n^{\bar{x}}, \mu\right) \leq 2m_n + 288r/n$. Consequently, the following holds for all n :

$$\mathrm{P}\left(T_{2,\rho}\left(L_n, \mu\right) \geq 2m_n + 288r/n\right) \leq be^{-ar}, \quad \forall r \geq 0$$

Reasoning as in the proof of Theorem (6.1.1) [194]:, one concludes that

$$T_{2,\rho}\left(\mathrm{V}, \mu\right) \leq \frac{288}{a} H\left(\mathrm{V}|\mu\right), \quad \forall \mathrm{V} \in \mathrm{P}\left(\mathbf{R}^d\right)$$

In this section, one considers more carefully the case $p=1$ of the preceding one. Let us recall that a probability measure μ on $\mathbf{R}^d$ satisfies the Poincare inequality with constant $C>0$ if

$$Var_\mu\left(f\right) \leq C\int\left|\nabla f\right|_2^2 d\mu \tag{11}$$

for all smooth f .

The following theorem proves the equivalence between Poincare inequality, dimension free exponential concentration and the corresponding transportation-cost inequality.

Theorem (6.1.11) [194]: Let μ be a probability measure on $\mathbf{R}^d$. The following propositions are equivalent:

(i) The probability measure μ verifies Poincare' inequality with a constant C_1

(ii) The probability measure μ verifies for some constants $a, b > 0$

$$\mu^n\left(A + D_{2,1}\left(r\right)\right) \geq 1 - be^{-ar}, \quad \forall r \geq 0$$

for all subset A of $\left(\mathbf{R}^d\right)^n$ such that $\mu^n\left(A\right) \geq 1/2$, where the set $D_{2,1}\left(r\right)$ is defined by

$$D_{2,1}\left(r\right) = \left\{x \in \left(\mathbf{R}^d\right)^n \ s.t. \sum_{i=1}^n \alpha_1\left(\left|x^i\right|_2\right) \leq r\right\}$$

(iii) The probability measure μ verifies the following transportation-cost inequality for some constant $C_2 > 0$.

$$T_{SG}(\mathrm{V},\mu)=\inf\int\alpha_1\left(|x-y|_2\right)d\pi(x,y)\le C_2H\left(\mathrm{V}|\mu\right),\quad \forall \mathrm{V}\in \mathrm{P}\left(\mathbf{R}^d\right)$$

More precisely:

- (i) implies (ii) with $a=K\max\left(C_{1,}\sqrt{C_1}\right)^{-1},K$ being a universal constant.

- (ii) implies (iii) with $C_2=2/a$.

- (iii) implies (i) with $C\ =C_2/2$.

The equivalence between (i) and (iii) was first obtained by Bobkov, Gentil and Ledoux in [256], with the Hamilton-Jacobi approach. The equivalence of (i) and (ii) or (ii) and (iii)) seems to be new.

Proof: Accoding to [255], (i) implies (ii) with $b=1$ and a depending only on C_1; one can take $a=K\max\left(C_{1,}\sqrt{C_1}\right)^{-1},K$, where K is a universal constant.

According to (a slightly different version of) Theorem (6.1.10) ,with $\rho=1$, (ii) implies (iii) (with $C_2=2/a$). It remains to prove that (iii) implies (i). This last point is classical; let us simply sketch the proof. The transportation-cost inequality is equivalent to the following property: for all bounded f on $\mathbf{R}^d$, $\int e^{Qf}d\mu\le e^{\int fd\mu}$ where $Qf(x)=\inf_{y\in\mathbf{R}^d}\left\{f(y)+C_2^{-1}\alpha_1\left(|x-y|_2\right)\right\}$ (see [254], [190]). Let f be a smooth function and apply the preceding inequality to tf . When t goes to 0, it can be shown that

$$Q(tf)(x)-tf(x)=-\frac{C_2t^2}{4}|\nabla f|_2^2(x)+0\left(t^2\right)$$

so $\int e^{Q(tf)}d\mu=1+t\int fd\mu+\frac{t^2}{2}\int f^2d\mu\frac{C_2t^2}{4}\int|\nabla f|_2^2d\mu+0\left(t^2\right)$. On the other hand,

$$e^t\int fd\mu=1+t\int fd\mu+\frac{t^2}{2}\int(fd\mu)^2$$

One concludes, that

$$Var(f)\le\frac{C_2}{2}\int|\nabla f|^2d\mu$$

which achieves the proof.

Transportation-cost inequalities are closely related to the so called (τ)

property introduced by Maurey in [17]. If $c(x,y)$ is a non negative function defined on some product space $\chi\times\chi$ and μ is a probability measure on χ, one says that (μ,c) has the (τ) property if for all non-negative f on χ,

$$\int e^{Qcf}\,d\mu\int e^{-f}\,d\mu\le 1$$

Where $Qcf(x)=\inf_{y\in\chi}\int\{f(y)+c(x,y)\}$. By Latala and Wojtaszczyk [138] provides an excellent introduction together with a lot of new results concerning this class of inequalities.

The (τ) property is in fact a sort of dual version of the transportation-cost inequality. This was first observed by Bobkov and Gotze in [254]. In the case of T_2, one can show that if μ verifies $T_2(C)$ then $\left(\mu,(2C)^{-1}|x-y|_2^2\right)$ has the (τ) property and conversely, if $\left(\mu,(2C)^{-1}|x-y|_2^2\right)$ has the (τ) property, then μ verifies $T_2(C)$. A general statement can be found in [189].

Several sufficient conditions for transportation-cost inequalities are known. Let us recall some of them. In [191], The author proved the following result:

Theorem (6.1.12) [194]: Let μ be a symmetric probability measure on $\mathbf{R}$ of the form $d\mu(x)=e^{-V(x)}dx$, with a smooth function such that $\lim_{x\to\infty}\frac{V''(x)}{V'(x)}=0$. Let $p\ge 1$, if V is such that $\lim_{x\to\infty}\frac{x^{p-1}}{V'(x)}<+\infty$, then μ verifies the transportation-cost inequality

$$\inf_{\pi\in\mathrm{P}(\mathrm{V},\mu)}\int\alpha_p(x-y)\,d\pi(x,y)\le CH(\mathrm{V}|\mu),\quad\forall\mathrm{V}\in\mathrm{P}(\mathbf{R})$$

Where $\alpha_p(u)=u^2$ if $|u|\le 1$ and $\alpha_p(u)=|u|^p$ if $|u|\ge 1$.

The case $p=2$ was first established by Cattiaux and Guillin in [204] with a completely different proof. Other cost functions α can be considered in place of the α_p. Furthermore, if μ satisfies Cheeger's inequality on $\mathbf{R}$, then a necessary and sufficient condition is known for the transportation-cost inequality associated to α (see [191]).

On $\mathbf{R}^d$, a relatively weak sufficient condition for T_2 (and other transportation-cost inequalities) was will be established by the author in [189]

Define $w^{(d)}:\mathbf{R}^d \to \mathbf{R}^d:(x_1,...,x_d)\mapsto(w(x_1),...,w(x_d))$, where $w(u)=\varepsilon(u)\max(|u|,u^2)$ with $\varepsilon(u)=1$ when u is non-negative and -1 otherwise. If the image of μ under the map $w^{(d)}$ verifies the Poincare inequality, then μ satisfies T_2. It can be shown that this condition is strictly weaker than the condition μ verifies LSI (see [189]).

Other sufficient conditions were obtained by Bobkov and Ledoux in [257] with an approach based on the Prekopa-Lcindler inequality, or in [33] by Cordero-Erausquin, Gangbo and Houdre with an optimal transportation method .

The following proposition is quite classical in Large Deviations theory. It can be found in DCllschdl and Strook's book [125].

Proposition (6.1.13) [194]: Let $A\subset \mathrm{P}(\chi)$ be such that $\{x\in\chi^n : L_n^x\in A\}$ is measurable. Then for every probability measure v on χ absolutely continuous with respect to μ and such that

$\mathrm{V}^n(x:L_n^x\in A)>0$, we have

$$\frac{1}{n}\log\left(\mu^n(L_n\in A)e^{nH(\mathrm{V}|\mu)}\right)\geq$$

$$-H(\mathrm{V}|\mu)\frac{\mathrm{V}^n(L_n\in A^c)}{\mathrm{V}^n(L_n\in A)}\frac{1}{n}\log\mathrm{V}^n(L_n\in A)\frac{1}{\mathrm{n}\ \ ne\mathrm{V}^n(L_n\in A)} \tag{12}$$

Proof: Let $h=\dfrac{dv^n}{d\mu^n}$ and $B=\{x\in\chi^n : L_n^x\in A \textit{ and } h(x)>0\}$. Then,

$$\mu^n(L_n\in A)\geq\mu^n(B)=\int_B h(x)d\mathrm{V}^n(x)=\mathrm{V}^n(B)\frac{\int_B e^{-\log h(x)}d\mathrm{V}^n(x)}{\mathrm{V}^n(B)}$$

Applying Jensen's inequality gives

$$\log\mu^n(L_n\in A)\geq\log\mathrm{V}^n(B)-\frac{\int_B\log h(x)d\mathrm{V}^n}{\mathrm{V}^n(B)}$$

Since $H(\mathrm{V}^n|\mu^n)=\int\log h(x)d\mathrm{V}^n$, one concludes that

$$\log\mu^n(L_n\in A)\geq\log\mathrm{V}^n(B)-\frac{H(\mathrm{V}^n|\mu^n)}{\mathrm{V}^n(B)}+\frac{\int_{B^c}\log h(x)h(x)d\mu^n}{\mathrm{V}^n(B)} \tag{13}$$

But for all $x > 0, \quad x\log x \geq -1/e$, so

$$\frac{\int_{B^c} \log h(x) h(x) d\mu^n}{\mathrm{V}^n(B)} \geq \frac{\mu^n(B)}{e\mathrm{V}^n(B)} \geq \frac{1}{e\mathrm{V}^n(B)} \tag{14}$$

Putting (14) into (13) and using $H(\mathrm{V}^n|\mu^n) = nH(\mathrm{V}|\mu)\ and \mathrm{V}^n(B) = \mathrm{V}^n(L_n \in A)$

gives the desired inequality.

Theorem (6.1.14) [194]: If $\mu \in \mathrm{P}_p(\chi)$, then for all t > 0,

$$\liminf_{n\to+\infty} \frac{1}{n} \log \mathrm{P}\left(\mathrm{W}_p(L_n,\mu) > t\right) \geq \inf\left\{H(\mathrm{V}|\mu) : \mathrm{V} \in \mathrm{P}_p(\chi)\, s.t.\, \mathrm{W}_p(\mathrm{V},\mu) > t\right\}$$

Proof: Let $t \geq 0$ and define $A = \left\{\mathrm{V} \in P_p(\chi)\, s.t.\, \mathrm{W}_p(\mathrm{V},\mu) > t\right\}$. Take $\mathrm{V} \in A$ such that $H(\mathrm{V}|\mu) < +\infty$. If $(Y_i)_i$ is an i.i.d sequence of law V, and $L_n^Y = n^{-1}\sum_{i=1}^n \delta_{Y_i}$, then L_n^Y converges to V almost surely for the W_ρ distance and so $\mathrm{V}^n(L_n \in A) = \mathrm{P}\left(\mathrm{W}_\rho(L_n^Y,\mu) > t\right) \to \mathrm{P}\left(\mathrm{W}_\rho(\mathrm{V},\mu) > t\right) = 1$, when n tends to $+\infty$. Applying Proposition (6.1.13) to A and V and taking the limit when n goes to $+\infty$, gives

$$\liminf_{n\to+\infty} \frac{1}{n} \log \mathrm{P}\left(\mathrm{W}_\rho(L_n,\mu) > t\right) \geq -H(\mathrm{V}|\mu).$$

Optimizing over V gives the result.

Corollary(6.1.15)[274]: If $\mu \in \mathrm{P}_p(\chi_k)$, then for all t > 0,

$$\liminf_{n\to+\infty} \frac{1}{n} \log \mathrm{P}\left(W_p(L_n,\mu) > t\right) \geq -\inf\left\{\sum_{j=1}^k H(v_j|\mu) : \sum_{j=1}^k v_j \in \mathrm{P}_p(\chi)\, s.t.\, \sum_{j=1}^k W_p(v_j,\mu) > t\right\}$$

Proof: Let $t \geq 0$ and define $A = \left\{\sum_{j=1}^k v_j \in \mathrm{P}_p(\chi_k)\, s.t.\, \sum_{j=1}^k W_p(v_j,\mu) > t\right\}$. Take $\sum_{j=1}^n v_j \in A$ such that $\sum_{j=1}^k H(v_j|\mu) < +\infty$. If $(Y_i)_i$ is an $i.i.d$ sequence of law $\sum_{j=1}^n v_j$, and $L_n^{\sum_{j=1}^k Y_j} = n^{-1}\sum_{j=1}^k \sum_{i=1}^n (\delta_{Y_i})_j$, then L_n^Y converges to $\sum_{j=1}^n v_j$ almost surely for the W_p distance and so

$\left(\sum_{j=1}^n v_j\right)^n (L_n \in A) = \mathrm{P}\left(W_p(L_n^Y,\mu) > t\right) \to \mathrm{P}\left(\sum_{j=1}^k W_p(v_j,\mu) > t\right) = 1$, when $n \longrightarrow +\infty$.

Applying [194] to A and $\sum_{j=1}^{k} v_j$ and taking the limit when $n \longrightarrow +\infty$, gives

$$\liminf_{n\to+\infty} \frac{1}{n} \log \mathrm{P}\left(W_p\left(L_n, \mu\right) > t\right) \geq -\sum_{j=1}^{k} H\left(v_j | \mu\right).$$

Optimizing over $\sum_{j=1}^{k} v_j$ gives the result.

Corollary(6.1.16)[274]: (Marton). If μ verifies $\mathrm{T}_1(\tilde{c})$, then for a measurable subset A of χ_k, such that $\mu(A) \geq \frac{1}{2}$.

$$\mu\left(A^{(\varepsilon+r_0)}\right) \geq 1 - be^{-\tilde{c}^{-1}\varepsilon^2}, \quad \forall \varepsilon \geq 0$$

where $r_0 = \sqrt{\tilde{c}\log(2)}$.

Proof: Consider a subset A of χ_k, and define $d\mu_A = \mathrm{I}_A d\mu\left(\sum_{j=1}^{k} x_j\right) / \mu(A)$. Let $B = \chi_k \setminus A^{(\varepsilon+r_0)}$ and define μ_B accordingly. Since the distance between two points of A and B is always more than $(\varepsilon + r_0)$, one has $W_1(\mu_A, \mu_B) \geq (\varepsilon + r_0)$. The triangle inequality and the transportation-cost inequality $\mathrm{T}_1(\tilde{c})$ yield

$$\begin{aligned}(\varepsilon + r_0) \leq W_1(\mu_A, \mu_B) &\leq W_1(\mu_A, \mu) + W_1(\mu_B, \mu) \\ &\leq \sqrt{\tilde{c}H\left(\mu_A | \mu\right)} + \sqrt{\tilde{c}H\left(\mu_B | \mu\right)} \\ &= \sqrt{\tilde{c}\log\left(1/\mu(A)\right)} + \sqrt{\tilde{c}\log\left(1/\mu(B)\right)}\end{aligned}$$

Rearranging terms gives the result.

Corollary (6.1.17)[274]: Let $\mu \in \mathrm{P}_2(\chi_k)$ and a > 0; the following Propositions are equivalent:

(i)There are $r_0, b \geq 0$ such that for all n the probability μ^n verifies (4),

(ii)The probability μ verifies $T_2(1/a)$.

Let us recall the definition of the series of the T_1 transportation-cost inequality. We say that a probability measure μ on χ_k verifies $\mathrm{T}_2(1/a)$, if

$$\sum_{j=1}^{k} W_1\left(v_j, \mu\right) \leq \sqrt{\tilde{c}\sum_{j=1}^{k} H\left(v_j | \mu\right)}, \text{for every} \sum_{j=1}^{k} v_j \in \mathrm{P}(\chi_k)$$

Proof: Let us show that (ii) implies (i). The main point is that T_2 tensorizes ; this means that if μ verifies $T_2(1/a)$ then μ^n verifies $T_2(1/a)$ on the space χ_k^n equipped with ρ_2^n (see[190]) .Jensen's inequality implies that $W_1^{\,2} \le T_2$ and consequently μ^n verifies $T_1(1/a)$(on χ_k^n equipped with ρ_2^n) for all n. Applying Proposition (6.1.3) to μ^n gives (1) with $r_0 = \sqrt{\log(2)/a}, \quad b = 1$ and a.

Let us show that (i) implies (ii). For every integer n, and $\sum_{j=1}^{k} x_j \in \chi_k^n$, define $L_n^{\left(\sum_{j=1}^{k} x_j\right)} = n^{-1} \sum_{j=1}^{k} \sum_{i=1}^{n} \delta_{\left(\sum_{j=1}^{k} x\right)^i}$. The map $\sum_{j=1}^{k} x_j \mapsto W_2\left(L_n^{\left(\sum_{j=1}^{k} x_j\right)}, \mu\right)$ is $\frac{1}{\sqrt{n}} - Lipschitz$.

Indeed, if

$$\sum_{j=1}^{k} x_j = \left(\left(\sum_{j=1}^{k} x_j\right)^1, \left(\sum_{j=1}^{k} x_j\right)^2, \ldots, \left(\sum_{j=1}^{k} x_j\right)^n\right) \text{and}$$

$$\sum_{j=1}^{k} y_j = \left(\left(\sum_{j=1}^{k} y_j\right)^1, \left(\sum_{j=1}^{k} y_j\right)^2, \ldots, \left(\sum_{j=1}^{k} y_j\right)^n\right)$$

are in χ_k^n, then thanks to the triangle inequality,

$$\left| W_2\left(L_n^{\left(\sum_{j=1}^{k} x_j\right)}, \mu\right) - W_2\left(L_n^{\left(\sum_{j=1}^{k} y_j\right)}, \mu\right) \right| \le W_2\left(L_n^{\left(\sum_{j=1}^{k} x_j\right)}, L_n^{\left(\sum_{j=1}^{k} y_j\right)}\right)$$

According to the convexity property of $T_2(.,.)$(see [37]), we have

$$T_2\left(L_n^{\left(\sum_{j=1}^{k} x_j\right)}, L_n^{\left(\sum_{j=1}^{k} y_j\right)}\right) \le \frac{1}{n} \sum_{i=1}^{n} T_2\left(\delta_{\left(\sum_{j=1}^{k} x_j\right)^i}, \delta_{\left(\sum_{j=1}^{k} y_j\right)^i}\right)$$

$$= \frac{1}{n} \sum_{i=1}^{n} \rho\left(\left(\sum_{j=1}^{k} x_j\right)^i, \left(\sum_{j=1}^{k} y_j\right)^i\right)^2 = \frac{1}{n} \rho_2^n \left(\sum_{j=1}^{k} x_j, \sum_{j=1}^{k} y_j\right)^2$$

which proves the claim.

Now, let $\left(\left(\sum_{j=1}^{k} X_j\right)_i\right)_i$ be an i.i.d sequence of law μ and let L_n be its empirical measure. Let m_n be the median of $W_2(L_n, \mu)$ and define

$A = \left\{ \sum_{j=1}^{k} x_j : W_2\left(L_n^{\left(\sum_{j=1}^{k} x_j\right)}, \mu \right) \le m_n \right\}$. Then $\mu^n(A) \ge 1/2$ and it is easy to show

that $A^{(\varepsilon + r_0)} \subset \left\{ \sum_{j=1}^{k} x_j : W_2\left(L_n^{\left(\sum_{j=1}^{k} x_j\right)}, \mu \right) \le m_n + \frac{(\varepsilon + r_0)}{\sqrt{n}} \right\}$.

Applying (4) to A gives

$$\mathrm{P}\left(W_2(L_n, \mu) > m_n + (\varepsilon + r_0)\sqrt{n} \right) \le b \ \exp\left(-a\varepsilon^2\right), \quad \forall \varepsilon \ge 0$$

Equivalently, as soon as $\sqrt{n}\left(\sum_{j-1}^{k} u_j - m_n \right) \ge r_0$, we have

$$\mathrm{P}\left(W_2(L_n, \mu) > \sum_{j-1}^{k} u_j \right) \le b \ \exp\left(-a\left(\sqrt{n}\left(\sum_{j-1}^{k} u_j - m_n \right) - r_0 \right)^2 \right).$$

Now, since $W_2(L_n, \mu)$ converges to 0 in probability (see Proposition (6.1.2)), the sequence $m_n \to 0$ when $n \to +\infty$. Consequently,

$$\limsup_{n \to +\infty} \frac{1}{n} \log \mathrm{P}\left(W_2(L_n, \mu) > \sum_{j=1}^{k} u_j \right) \le -a\left(\sum_{j-1}^{k} u_j \right)^2, \quad \forall \sum_{j-1}^{k} u_j > 0$$

The final step is given by Large deviations. According to Theorem(6.1.4),

$$\liminf_{n \to +\infty} \frac{1}{n} \log \mathrm{P}\left(W_2(L_n, \mu) > \sum_{j-1}^{k} u_j \right) \ge -\inf\left\{ \sum_{j=1}^{k} H\left(v_j | \mu\right) : \sum_{j=1}^{k} v_j \in \mathrm{P}_2(\chi_k) \, s.t. \sum_{j=1}^{k} W_2\left(v_j, \mu\right) > \sum_{j=1}^{k} u_j \right\}$$

This together with the preceding inequality yields

$$\inf\left\{ \sum_{j=1}^{k} H\left(v_j | \mu\right) : \sum_{j=1}^{k} v_j \in \mathrm{P}_2(\chi_k) \, s.t. \sum_{j=1}^{k} W_2\left(v_j, \mu\right) > \sum_{j=1}^{k} u_j \right\} \ge a\left(\sum_{j=1}^{k} u_j \right)^2$$

or in other words,

$$a\sum_{j=1}^{k} W_2\left(v_j, \mu\right)^2 \le \sum_{j=1}^{k} H\left(v_j | \mu\right)$$

and this achieves the proof.

Corollary(6.1.18)[274]: Let μ , be a probability measure on χ_k and suppose that for all integer n the function F_n defined on χ_k^n by

$F_n\left(\sum_{j=1}^{k} x_j\right) = W_2\left(L_n^{\sum_{j=1}^{k} x_j}, \mu\right)$ verifies

$$\sum_{i=1}^{n} \left|\nabla_i F_n\right|^2 \left(\sum_{j=1}^{k} x_j\right) \le \frac{1}{n} \text{ for } \mu^n \text{ almost every } \sum_{j=1}^{k} x_j \in \chi_k^n \qquad (15)$$

If μ , verifies the inequality $\mathrm{LSI}(\tilde{c})$, then μ , verifies the inequality $\mathrm{T}_2(\tilde{c})$. Suppose that $\chi_k = \mathbf{R}^d$ or a Riemannian manifold M, then according to Rademacher's Theorem, F_n is almost everywhere differentiable on $\left(\mathbf{R}^d\right)^n$ (resp. M^n) with respect to the Lebesgue measure. It is thus easy to show that condition (15) is fulfilled when μ , is absolutely continuous with respect to Lebesgue measure. This permits us to recover Otto and Villani's result as stated in [32].

Proof: As we said above the product measure μ^n verifies the inequality (6). Apply this inequality to $f = e^{\frac{s}{2}F_n}$, with $s \in \mathbf{R}^+$. It is easy to show that $\left|\nabla_i e^{\frac{s}{2}F}\right| = \frac{s}{2} e^{\frac{s}{2}F_n} \left|\nabla_i F_n\right|$, thus, using condition (15), we see that the right hand side of (6) is less than $\tilde{c}\frac{s^2}{4n}\int e^{sF_n} d\mu^n$.

Letting $\mathrm{Z}(s) = \int e^{sF_n} d\mu^n$, we get the differential inequality:

$$\frac{\mathrm{Z}'(s)}{s\,\mathrm{Z}(s)} - \frac{\log \mathrm{Z}(s)}{s^2} \le \frac{\tilde{c}}{4n}$$

Integrating this yields:

$$\mathrm{Z}(s) = \int e^{sF_n} d\mu^n \le e^{s\int F_n d\mu^n + \frac{\tilde{c}s^2}{4n}}, \quad \forall s \in \mathbf{R}^+$$

This implies that

$$\mathrm{P}\left(W_2\left(L_n, \mu\right) \ge t + \mathrm{E}\left[W_2\left(L_n, \mu\right)\right]\right) \le e^{-nt^2/\tilde{c}}$$

Corollary(6.1.19)[274]: If μ verifies the inequality $\mathrm{LSI}^+(\tilde{c})$, then μ verifies $\mathrm{T}_2(\tilde{c})$.

.**Proof:** The inequality LSI^+ tensorizes, so μ^n verifies[194].

$$H\left(\mu^n \left| f^2\right.\right) \le c\int \sum_{i=1}^{n} \left|\nabla_i^- f\right|^2 d\mu^n$$

Take $f = e^{\frac{s}{2}F_n}$, $s \in \mathbf{R}^+$ with $F_n\left(\sum_{j=1}^{k} x_j\right) = W_2\left(L_n^{\left(\sum_{j=1}^{k} x_j\right)}, \mu\right)$. Once again, it is easy to check that $\left|\nabla_i^- e^{\frac{s}{2}F_n}\right| = \frac{s}{2} e^{\frac{s}{2}F_n}\left|\nabla_i^- F_n\right|$ (note that the function $\sum_{j=1}^{k} x_j \mapsto e^{s\left(\sum_{j=1}^{k} x_j\right)}$ is non decreasing). Reasoning as in the proof of Theorem (6.1.7) it is enough to show that $\sum_i \left|\nabla_i^- F_n\right|^2 \left(\sum_{j=1}^{k} x_j\right) \le 1/n$ for μ^{n-} almost all $\sum_{j=1}^{k} x_j \in \chi_k^n$. Let us show how to compute $\left|\nabla_i^- F_n\right|$. Let $\sum_{j=1}^{k} z_j \in \sum_{j=1}^{k} X_j$, $\sum_{j=1}^{k} a_j = \left(\left(\sum_{j=1}^{k} a_j\right)^1, \ldots, \left(\sum_{j=1}^{k} a_j\right)^n\right) \in \chi_k^n$ and set

$$\sum_{j=1}^{k} z_j a_j = \left(\sum_{j=1}^{k} z_j, \left(\sum_{j=1}^{k} a_j\right)^2, \ldots, \left(\sum_{j=1}^{k} a_j\right)^n\right)$$

$$\left|\nabla_i^- F_n\right|\left(\sum_{j=1}^{k} a_j\right) = \frac{1}{2F_n\left(\left(\sum_{j=1}^{k} a_j\right)\right)} \limsup_{z \to a^1} \frac{\left[\mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} z_j a_j\right)}, \mu\right) - \mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)}, \mu\right)\right]_+}{\rho\left(\sum_{j=1}^{k} a_j, \left(\sum_{j=1}^{k} a_j\right)^1\right)}$$

Let $\pi \in \mathrm{P}\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)}, \mu\right)$ be an optimal coupling; it is not difficult to see that we can write

$\pi\left(d\left(\sum_{j=1}^{k} x_j\right), d\left(\sum_{j=1}^{k} y_j\right)\right) = p\left(\left(\sum_{j=1}^{k} x_j\right), d\left(\sum_{j=1}^{k} y_j\right)\right) L_n^{\left(\sum_{j=1}^{k} a_j\right)}\left(d\left(\sum_{j=1}^{k} x_j\right)\right)$, where $p\left(\left(\sum_{j=1}^{k} a_j\right)^i, d\sum_{j=1}^{k} y_j\right) = \left(\sum_{j=1}^{k} v_j\right)_i \left(d\left(\sum_{j=1}^{k} y_j\right)\right)$ with $\left(\sum_{j=1}^{k} v_j\right)_1, \ldots, \left(\sum_{j=1}^{k} v_j\right)_n$ probability measures on χ_k such that $n^{-1}\left(\left(\sum_{j=1}^{k} v_j\right)_1 + \ldots + \left(\sum_{j=1}^{k} v_j\right)_n\right) = \mu$.

Let $\tilde{p}$ be defined as p with $\left(\sum_{j=1}^{k} z_j\right)$ in place of $\left(\sum_{j=1}^{k} v_j\right)^1$; then

$$\tilde{\pi} = \tilde{p}\left(\sum_{j=1}^{k} x_j, d\left(\sum_{j=1}^{k} y_j\right)\right) L_n^{\left(\sum_{j=1}^{k} z_j a_j\right)}\left(d\left(\sum_{j=1}^{k} y_j\right)\right) \text{ belongs to } P\left(L_n^{\left(\sum_{j=1}^{k} z_j a_j\right)}, \mu\right)$$

We have

$$\mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} z_j a_j\right)}, \mu\right) - \mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)}, \mu\right) \le \int \rho\left(\sum_{j=1}^{k} x_j, \sum_{j=1}^{k} y_j\right)^2 d\tilde{\pi}\left(\sum_{j=1}^{k} x_j, \sum_{j=1}^{k} y_j\right) - \int \rho\left(\sum_{j=1}^{k} x_j, \sum_{j=1}^{k} y_j\right)^2 d\pi\left(\sum_{j=1}^{k} x_j, \sum_{j=1}^{k} y_j\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\int \rho\left(\left(\sum_{j=1}^{k} z_j a_j\right)^i, \sum_{j=1}^{k} y_j\right)^2 d\left(\sum_{j=1}^{k} v_j\right)_i \left(\sum_{j=1}^{k} y_j\right) - \frac{1}{n}\sum_{i=1}^{n}\int \rho\left(\left(\sum_{j=1}^{k} a_j\right)^i, \sum_{j=1}^{k} y_j\right)^2 d\left(\sum_{j=1}^{k} v_j\right)_i \left(\sum_{j=1}^{k} y_j\right)$$

$$= \frac{1}{n}\int \rho\left(\sum_{j=1}^{k} z_j, \sum_{j=1}^{k} y_j\right)^2 - \rho\left(\left(\sum_{j=1}^{k} a_j\right)^1, \sum_{j=1}^{k} y_j\right)^2 d\left(\sum_{j=1}^{k} v_j\right)_1 \left(\sum_{j=1}^{k} y_j\right)$$

$$\le \frac{1}{n}\rho\left(\sum_{j=1}^{k} z_j, \left(\sum_{j=1}^{k} a_j\right)^1\right)\int \rho\left(\sum_{j=1}^{k} z_j, \sum_{j=1}^{k} y_j\right) + \rho\left(\left(\sum_{j=1}^{k} a_j\right)^1, \sum_{j=1}^{k} y_j\right) d\left(\sum_{j=1}^{k} v_j\right)_1 \left(\sum_{j=1}^{k} y_j\right)$$

Since the function $\sum_{j=1}^{k} x_j \mapsto \left[\sum_{j=1}^{k} x_j\right]_+$ is non decreasing, we have

$$\frac{\left[\mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} z_j a_j\right)}, \mu\right) - \mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)}, \mu\right)\right]_+}{\rho\left(\sum_{j=1}^{k} z_j, \left(\sum_{j=1}^{k} a_j\right)^1\right)} \le \frac{1}{n}\int \rho\left(\sum_{j=1}^{k} z_j, \sum_{j=1}^{k} y_j\right) + \rho\left(\left(\sum_{j=1}^{k} a_j\right)^1, \sum_{j=1}^{k} y_j\right) d\left(\sum_{j=1}^{k} v_j\right)_1 \left(\sum_{j=1}^{k} y_j\right)$$

Letting $\sum_{j=1}^{k} z_j \to \left(\sum_{j=1}^{k} a_j\right)^1$ yields

$$\left|\nabla_i^- F_n\right|\left(\sum_{j=1}^{k} a_j\right)^2 \le \frac{\int \rho\left(\left(\sum_{j=1}^{k} a_j\right)^1, \sum_{j=1}^{k} a_j\right)^2 d\left(\sum_{j=1}^{k} v_j\right)_1 \left(\sum_{j=1}^{k} y_j\right)}{n^2 \mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)}, \mu\right)}.$$ Doing the same

computations for the other derivatives (with the same optimal coupling π), we get

$$\left|\nabla_i^- F_n\right|\left(\sum_{j=1}^{k} a_j\right)^2 \le \frac{\int \rho\left(\left(\sum_{j=1}^{k} a_j\right)^i, \sum_{j=1}^{k} y_j\right)^2 d\left(\sum_{j=1}^{k} v_j\right)_i \left(\sum_{j=1}^{k} y_j\right)}{n^2 \mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)}, \mu\right)}.$$

Summing these inequalities gives $\sum_i \left|\nabla_i^- F_n\right|^2 \left(\sum_{j=1}^{k} a_j\right) \le 1/n$ for all $\sum_{j=1}^{k} a_j \in \chi_k^n$, which achieves the proof.

Corollary(6.1.20)[274]: Let μ be a probability measure on $\mathbf{R}^d$ and $p \in [1,2]$. The following propositions are equivalent:

(i)The two level concentration (10) holds for some non-negative $\sum_{j=1}^{k} a_j, \sum_{j=1}^{k} b_j$ independent of n

(ii)The probability measure μ verifies the series of transportation-cost inequality

$$\sum_{j=1}^{k} \mathrm{T}_{2,p}\left(v_j, \mu\right) \le \tilde{c} \sum_{j=1}^{k} H\left(v_j \,|\mu\right), \quad \forall \sum_{j=1}^{n} v_j \in \mathrm{P}\left(\mathbf{R}^d\right)$$

for some constant $\tilde{c}$.

More precisely, if (10) holds for some constants $\sum_{j=1}^{k} a_j, \sum_{j=1}^{k} b_j$, then the series of the series of the transportation-cost inequality holds with the constant $\tilde{c} = \frac{\tilde{c}}{\sum_{j=1}^{k} a_j}$ (for $j = 1, \tilde{c} = 288$ see[194]). Conversely, if the transportation-cost inequality holds for some constant $\tilde{c}$, then (10) is true for $\sum_{j=1}^{k} b_j = 2$ and

$$\sum_{j=1}^{k} a_j = 1/(2\tilde{c})$$

Proof: Let us recall the proof of (ii) implies (i). According to the tensorization property, for all n and all probability measure v on $\left(\mathbf{R}^d\right)^n$,

$$T_{2,p}\left(\sum_{j=1}^{n} v_j , \mu^n\right) \le \tilde{c}H\left(\sum_{j=1}^{n} v_j \,\middle|\, \mu^n\right)$$

holds. Take A and B in $\left(\mathbf{R}^d\right)^n$ and define $d\mu_A^n = \mathrm{I}_A d\mu / \mu^n(A)$ and $d\mu_B^n = \mathrm{I}_B d\mu / \mu^n(B)$. According to point (ii) of Lemma(6.1.9) and the transportation-cost inequality satisfied by μ^n, one has

$$\begin{aligned} T_{2,p}\left(\mu_A^n, \mu_B^n\right) &\le 2T_{2,p}\left(\mu_A^n, \mu^n\right) + 2T_{2,p}\left(\mu_B^n, \mu^n\right) \\ &\le 2\tilde{c}H\left(\mu_A^n \,\middle|\, \mu^n\right) + 2\tilde{c}H\left(\mu_B^n \,\middle|\, \mu^n\right) \\ &= -2\tilde{c}\log\left(\mu^n(A)\mu^n(B)\right) \end{aligned}$$

Define

$$\tilde{c}_{2,p}(A,B) = \inf\left\{(\varepsilon + r_0) \ge 0 \; s.t. \; \left(A + B_{2,p}(\varepsilon + r_0)\right) \cap B \ne \phi\right\}$$

then $T_{2,p}\left(\mu_A^n, \mu_B^n\right) \ge \tilde{c}_{2,p}(A,B)$ and so

$$\mu^n(A)\mu^n(B) \le e^{-\tilde{c}_{2,p}(A,B)/2\tilde{C}}.$$

Now, if $\mu^n(A) \ge \dfrac{1}{2}$ and $B = \left(\mathbf{R}^d\right)^n \setminus \left(A + B_{2.p}(\varepsilon + r_0)\right)$, one has $c_{2,p}(A,B) = (\varepsilon + r_0)$ and so $\mu^n\left(A + B_{2,p}(\varepsilon + r_0)\right) \ge 1 - 2e^{-(\varepsilon + r_0)/2\tilde{c}}$. Using point (iii) of Lemma (6.1.9) gives

$$\mu^n\left(A + \sqrt{(\varepsilon + r_0)}B_2 + \sqrt[p]{(\varepsilon + r_0)}B_p\right) \ge 1 - 2e^{-(\varepsilon + r_0)/2\tilde{c}}.$$

Now let us prove the converse. Let $(X_i)_i$ be an i.i.d sequence of law μ and let L_n be its empirical measure. Consider $A = \left\{\sum_{j=1}^{k} x_j \in \left(\mathbf{R}^d\right)^n \; s.t. \; T_{2,p}\left(L_n^{\left(\sum_{j=1}^{k} x_j\right)}, \mu\right) \le m_n\right\}$ where m_n denotes the median of $T_{2,p}\left(L_n^{\left(\sum_{j=1}^{k} x_j\right)}, \mu\right)$. According to point (iii) of Lemma(6.1.9)

$A + \sqrt{(\varepsilon + r_0)}B_2 + \sqrt[p]{(\varepsilon + r_0)}B_p \subset A + 12B_{2,p}(\varepsilon + r_0)$. Let $\sum_{j=1}^{k} x_j \in A + 12B_{2,p}(\varepsilon + r_0)$; there is some $\sum_{j=1}^{k} \bar{x}_j \in A + 12B_{2,p}(\varepsilon + r_0)$ there is some $\sum_{j=1}^{k} \bar{x}_j \in A$ such that

$$\sum_{i=1}^{n}\sum_{j=1}^{d}\alpha_p\left(\frac{x_j^i-\overline{x}_j^i}{12}\right)\le(\varepsilon+r_0)$$

(here $\sum_{j=1}^{k}x_j=\left(\left(\sum_{j=1}^{k}x_j\right)^1,\left(\sum_{j=1}^{k}x_j\right)^2,...,\left(\sum_{j=1}^{k}x_j\right)^n\right)$ with $\left(\sum_{j=1}^{k}x_j\right)^i\in\mathbf{R}^d$). Since

$\frac{\alpha_p\left(\sum_{j=1}^{k}x_j/12\right)\ge\alpha_p\left(\sum_{j=1}^{k}x_j\right)}{\tilde{c}_1}$, we get $\mathrm{T}_{2,p}\left(L_n^{\left(\sum_{j=1}^{k}x_j\right)},L_n^{\left(\sum_{j=1}^{k}\overline{x}_j\right)}\right)\le\frac{\tilde{c}_1(\varepsilon+r_0)}{n}$.

(for $j=1,\tilde{c}_3=144$ (see [194]). According to point (ii) of Lemma(6.1.9)

$$\mathrm{T}_{2,p}\left(L_n^{\left(\sum_{j=1}^{k}x_j\right)},\mu\right)\le2\mathrm{T}_{2,p}\left(L_n^{\left(\sum_{j=1}^{k}x_j\right)},L_n^{\left(\sum_{j=1}^{k}\overline{x}_j\right)}\right)+2\mathrm{T}_{2,p}\left(L_n^{\left(\sum_{j=1}^{k}\overline{x}_j\right)},\mu\right)\le2m_n+\frac{\tilde{c}(\varepsilon+r_0)}{n}.$$

Consequently, the following holds for all n :

$$\mathrm{P}\left(\mathrm{T}_{2,p}(L_n,\mu)\ge2m_n+\tilde{c}(\varepsilon+r_0)/n\right)\le\sum_{j=1}^{k}b_je^{-\sum_{j=1}^{k}a_j(\varepsilon+r_0)},\quad\forall\varepsilon\ge-r_0$$

Reasoning as in the proof of Theorem(6.1.6) we conclude that

$$\sum_{j=1}^{n}\mathrm{T}_{2,p}(v_j,\mu)\le\frac{\tilde{c}}{\sum_{j=1}^{k}a_j}\sum_{j=1}^{n}H(v_j|\mu),\ \text{for every}\ \sum_{j=1}^{n}v_j\in\mathrm{P}\left(\mathbf{R}^d\right)$$

Remark[6.1.21]: (i) If $\mu^n(A)=\frac{1}{2}$ we have $\frac{1}{2}\mu^n(B)\le e^{-\tilde{c}_{2,p}(A,B)/2\tilde{c}}$,approximately ,we have for $\tilde{c}_{2,p}(A,B)=0$ that $\mu^n(B)\le2$

(ii) We can deduce that $2e^{-\tilde{c}_{2,p}(A,B)/2\tilde{c}}\ge1$ and $\tilde{c}_{2,p}(A,B)\le2\log2$ .Hence $\tilde{c}\ge\frac{(\varepsilon+r_0)}{2\log2}$

Section (6.2): Poincare'Inequalities and Dimension of freeConcentration of Measure.

We say that a probability measure on a metric space (X,d) satisfies a Poincar´e inequality also called spectral gap inequality with the constant C, if for all locally Lipschitz function f, we have

$$Var(f) \leq C\int |\nabla f|^2 d\mu, \tag{16}$$

where the length of the gradient is defined by

$$|\nabla f|(x) := \limsup_{y \to x} \frac{|f(x) - f(y)|}{d(x,y)} \tag{17}$$

(when x is not an accumulation point of X, one defines $|\nabla f|(x) = 0$.

It is well known since the works [165], [244],[243] and [255] that the inequality (16) implies dimension free concentration inequalities for the product measures $\mu^n, n \geq 1$.

For example, in [255] M. Ledoux and S.G. Bobkov proved that if μ verifies (16), then there exists a constant L depending only on C such that for all subset A of X^n with $\mu^n(A) \geq \frac{1}{2}$,

$$\mu^n\left(A^h\right) \geq 1 - e^{-lh}, \forall h \geq 0 \tag{18}$$

where the set A^h is the enlargement of A defined by

$$A^h = \left\{ y \in X^n \inf_{x \in A} \sum_{i=1}^{n} \alpha\left(d\left(x_i, y_i\right)\right) \leq h \right\},$$

where $\alpha(u) = \min\left(|u|, u^2\right)$ for all $u \in R$ (see [255]

Inequalities such as (18) were first obtained by M. Talagrand in different articles using completely different techniques (see [181]).

In this paper, one will say that a probability measure μ satisfies the classical Poincar´e inequality with constant $C > 0$ on $\mathbf{R}^d$, if μ satisfies (16) on $\mathbf{R}^d$ equipped with its standard Euclidean norm $|\cdot|_2$. In that case, one will write that μ satisfies the inequality $\mathrm{SG}(C)$, where SG stands for spectral gap. In all the sequel, Bp will denote the ℓ^p unit ball of $\mathbf{R}^m$:

$$B_p = \left\{ x \in \mathbf{R}^m : |x_1|^p + \ldots + |x_m|^p \leq 1 \right\}$$

If μ satisfies the inequality $\mathrm{SG}(C)$ on $\mathbf{R}^d$ then (18) can be rewritten in a more pleasant way: for all subset A of $\left(\mathbf{R}^d\right)^n$ with $\mu^n(A) \geq \frac{1}{2}$,

$$\mu^n\left(A + \sqrt{h}B_2 + hB_1\right) \geq 1 - e^{-hl} \forall h \geq 0 \tag{19}$$

with a constant L depending on C and the dimension d. The archetypic example of a measure satisfying the classical Poincare' inequality is the exponential measure on R^d V_1^d where $dv_1(x)=\frac{1}{2}e^{-|x|}dx$. For this probability, (19) cannot be improved (a version of (19) with sharp constants has been established by Talagrand in [180] see also Maurey [17] Thus (19) expresses that the probability measures μ^n concentrate at least as fast as the exponential measure on $(\mathbf{R}^d)^n$.

Some probability measures concentrate faster than the exponential measure. For example, the standard Gaussian measure μ^n on $\mathbf{R}^m$ verifies for all $A\subset\mathbf{R}^m$ with $\mu^n(A)\geq\frac{1}{2}$

$$\mu^n\left(A+hB_2\right)\geq 1-e^{-h^2/2} \tag{20}$$

One cannot derive such a bound from the classical Poincar´e inequality. The inequality (20) requires stronger tools. For example, it is now well known that (20) follows from the Logarithmic-Sobolev inequality, introduced by L. Gross in [155], which is strictly stronger than the classical Poincare' inequality (see [174]). Let us recall, that a probability measure μ on $\mathbf{R}^d$ is said to satisfy the Logarithmic-Sobolev inequality with a constant $C>0$, if

$$H\left(\mu\middle|f^2\right)\leq C\int|\nabla f|_2^2\,d\mu \tag{21}$$

holds for all locally Lipschitz function f on $\mathbf{R}^d$, where the entropy functional is defined by

$$H_\mu(p)=\int f\log(f)d\mu-\left(\int fd\mu\right).\log\left(\int fd\mu\right),\forall f\geq 0$$

The aim is to show that considering Poincare´ inequality on R^d equipped with other metrics than the Euclidean distance makes possible to reach a large scope of concentration properties including Gaussian or even stronger behaviors. The metrics we are going to equip R^d with are of the form:

$$\mathrm{W}(x,y)=\left[\sum_{i=1}^{d}\left|w(x_i)-w(y_i)\right|^2\right]^{\frac{1}{2}},\forall x,y\in\mathbf{R}^d \tag{22}$$

We will assume that $\omega:\mathbf{R}\to\mathbf{R}$ is increasing and verifies:

(i) w is such that $x\mapsto w(x)/x$ is nondecreasing on $(0+\infty)$,(ii) w is non negative on $\mathbf{R}^+$,

(iii) w is such that $w(-x) = -w(x)$, for all $x \in \mathbf{R}$.

Note that the first assumption is verified as soon as w is convex on $\mathbf{R}^+$ with $w(0) = 0$.

Definition (6.2.1) [189]: One says that a probability measure μ on $\mathbf{R}^+$ satisfies the inequality $\mathrm{SG}(w,C)$ if μ satisfies the Poincar´e inequality (16) for the distance $\mathrm{W}(.,.)$ defined by (22) with the constant $C > 0$.

The following proposition gives examples of the variety of concentration rates enabled by our approach.

This result will be easily deduced from (18) and from an elementary comparison between the metric $\mathrm{W}_p(.,.)$ and the norms $|.|_p$.

This section will provide a lot of sufficient conditions for the inequalities $\mathrm{SG}(w,C)$. Let us just say for the moment that, in particular, for all $p \in [1,+\infty)$, the probability measure $dV_p(x) = \frac{1}{Z^p} e^{-|x|^p} dx$ verifies $\mathrm{SG}(w_p, C)$ for some C on $\mathbf{R}$. For these V_p one thus formally recovers a famous result by Talagrand [182]. Let us emphasize here that the above proposition only gives an example of the concentration results we can obtain with this approach. It is for instance possible to derive adapted concentration results for fast decreasing probabilities such as $d\mu(x) = \frac{1}{Z} \exp\left(-\exp\left(x^2\right)\right) dx$

Before presenting in details our results, let us outline some of the positive features of the inequalities $\mathrm{SG}(w,.)$:

(i) They enjoy the classical properties of Poincar´e inequalities: tensorization and stability under bounded perturbation.

(ii) A lot of workable sufficient conditions are available. In dimension one, one proves a necessary and sufficient condition.

(iii) A large variety of Talagrand's like concentration inequalities can be obtained. Moreover it is interesting to note that the same family of functional inequalities yields as well subgaussian and supergaussian estimates.

(iv) These inequalities are weak. For example, we are going to show that for all $p \in [1.2]$ the Poincar´e inequality $\mathrm{SG}(w_p,.)$ is strictly weaker than the Lata la-Oleszkiewicz inequality $\mathrm{LO}(p,.)$ defined below and gives the same kind of concentration.

(v) Finally, inequalities $\mathrm{SG}(w,.)$ are equivalent to certain transportation-cost nequalities and inf-convolution inequalities. As a byproduct, our section furnishes new results for these inequalities.

Let $p \in [1.2]$, one will say that a probability measure μ on $\mathbf{R}^d$ satisfies the inequality $\mathrm{LO}(P,C)$ if

$$\sup_{a\in(1,2)} \frac{\int f^2 d\mu - \left(\int |f|^a \, d\mu\right)}{(2-a)^{2\left(1-\frac{1}{p}\right)}} \le C\int |\nabla f|_2^2 \, d\mu \tag{23}$$

holds for all f smooth enough. For $p=1$, the inequality (23) is Poincar´e inequality $\mathrm{SG}(C)$ and for $p=2$ it is equivalent to the Logarithmic-Sobolev inequality see [239,].The $\mathrm{LO}(P,C)$ inequalities on $\mathbf{R}$ were completely characterized by Barthe and Roberto in [58]. Several extensions of this inequality were considered (see [73] or [59]). According to [239,], if μ is a probability measure on $\mathbf{R}^d$ satisfying $\mathrm{LO}(P,C)$, then there is a constant $\mathrm{L}>0$ such that μ^n verifies the concentration inequality (23). So, roughly speaking, if μ verifies $\mathrm{LO}(P,C)$ it concentrates independently of the dimension like $dv_p(x) = \frac{1}{Z_p} e^{-|x|^p} dx, p \in [1,2]$

These inequalities first appear in a paper of S. G. Bobkov and M. Ledoux [200]. Let H : $R \to R^+$ be a convex function ; one says that a probability μ on $\mathbf{R}^d$ verifies the modified Logarithmic-Sobolev inequality $\mathrm{LS}(H_q, C)$, if

$$H_\mu(f^2) \le C\int \sum_{i=1}^d H\left(\frac{\partial_i f}{f}\right) f^2 d\mu \tag{24}$$

holds for all positive and locally Lipschitz function f. When $H(x)=x^2$, the preceding inequality is simply the Logarithmic-Sobolev inequality, and if $H(x)=x^2$ for $|x|\le 1$ and ∞ otherwise, the resulting inequality was shown to be equivalent to the Poincar'e inequality (see [255]).

Many different tools are considered in order to obtain dimension free concentration estimates such as (23) and (24) for $1<p\le 2$ (see [176], [179], [239], [58], [59], [101], [57]) and p > 2 ([257], [75], [252], [33], [103], [193]). It will be a difficult task to give a complete summary of these various attempts. We will focus on four important functional approaches to the concentration of measure phenomenon: the Lata la-Oleszkiewicz inequalities, the modified logarithmic Sobolev inequalities, the super Poincar´e inequalities and the transportation-cost inequalities.

(i) The Lata la-Oleszkiewicz inequalities.We have already indicate how the concentration inequalities (23) for $\rho=1$ and $\rho=2$ can be derived from the classical Poincar´e inequality and the Logarithmic Sobolev inequality (21) respectively. In [239], R. Lata la and K. Oleszkiewicz proposed a family of inequalities interpolating between Poincar´e and Log-Sobolev. These inequalities are defined as follows.

Let $P\geq 2$ and consider $H_2(x)=|x|^q$ with $\frac{1}{p}+\frac{1}{q}=1$; the inequality $\mathrm{LS}(H_q,C)$ was studied by S. G. Bobkov and M. Ledoux in [257] and by S. G. Bobkov and B. Zegarlinski in [252], where a complete characterization on R was achieved (see [252]). This inequality is associated to supergaussian concentration. More precisely, if μ verifies $\mathrm{LS}(H_q,C)$ then for all subset A of $(\mathbf{R}^d)^n$ with $\mu^n(A)\geq\frac{1}{2}$,

$$\mu^n\left(A+t^{\frac{1}{p}}B_p\right)\geq 1-e^{-lt},\forall t\geq 0$$

where L is independent of n. For $p\geq 2$, the measure $dv_p(x)=\frac{1}{Z_p}e^{-|x|^p}dx$

verifies $\mathrm{LSI}(H_q,C)$ for some C and $\frac{1}{p}+\frac{1}{q}=1$.

Let $p\in[1,2]$ and consider $H_q(x)=\max\left(x^2,|x|^q\right)$ with $\frac{1}{p}+\frac{1}{q}=1$.

The family $\mathrm{LSI}(H_q,C)$ was first studied by 1 in [101] where it was shown that $\mathrm{LSI}(H_q,C)$ was fulfilled by $dv_p(x)=\frac{1}{Z_p}e^{-|x|^p}dx$ for $p\in[1,2]$ and $\frac{1}{p}+\frac{1}{q}=1$. It was recently completely characterized on the real line by F. Barthe and C. Roberto (see [59]). As shown in [101] or [59], if μ verifies $\mathrm{LS}(H_q,C)$ for some C then it verifies the concentration inequality (18) for some $L>0$. Other choices of H were considered in [59] and a general concentration inequality established (see [59,]). These results are available under the assumption that $H(x)/x^2$ is increasing. The resulting concentrations inequalities are thus always subgaussian.

Let $\beta:[1,+\infty)\to R^{+}$ be a nonincreasing function; one says that a probability μ on $\mathbf{R}^{d}$ verifies the super Poincar´e inequality with the function β if

$$\int f^{2}d\mu \le \beta(s)\int \left|\nabla f\right|_{2}^{2}d\mu + s\left(\int |f|\,d\mu\right)^{2}, \forall s\ge 1 \tag{25}$$

holds true for all locally Lipschitz function f. If μ verifies (25), one will write for short that μ satisfies the inequality $\mathrm{SP}(\beta)$. Super Poincare' inequalities were introduced by F. Y. Wang in [75]. They are of great interest in spectral theory or for isoperimetric problems (see [60]). Another nice feature of this family is that several other functional inequalities are encoded among it, i.e correspond to specific choices of (β). For example, defining for all $p\ge 1, \beta_{p}(s)=\left(\log(e+s)^{-2(1-1/p)}\right)$ then the Lata la-Oleszkiewicz inequality $\mathrm{LO}(P,C)$, $p\in[1,2]$ is equivalent to $\mathrm{SP}\left(\tilde{C}\beta_{p}\right)$ for some $\tilde{C}$ as shown in [73]. The same is true for F-Sobolev inequalities (see [75]) or Weak Logarithmic Sobolev inequalities (see [205]). For a general (β) only quite rough concentration estimates can be deduced from $\mathrm{SP}(\beta)$. For example, if μ verifies the inequality $SP\left(\tilde{C}\beta_{p}\right)$ for some C with the function $\left(\beta_{p}\right)$ defined above, then $\int e^{a|x|^{p}}d\mu(x)<+\infty$ for some $a>0$. The general case is more intricate (see [75] the present section). Moreover, unlike the functional inequalities presented above, the super Poincar´e inequality does not tensorize properly and thus the concentration bounds may be affected by the dimension.Transportation-cost inequalities were first introduced by K. Marton and M. Talagrand in [135, 139] and [79]. In these inequalities one tries to bound an optimal transportation-cost in the sense of Kantorovich by the relative entropy functional.

More precisely, if $c:X\times X\to \mathbf{R}^{+}$ is a measurable map on some metric space X, the optimal transportation-cost between V and μ $\in \mathrm{P}(X)$(the set of probability measures on X) is defined by

$$T_{c}(\mathrm{V},\mu)=\inf_{\pi\in p(v,\mu)}\int c(x,y)\,d\pi \tag{26}$$

where $\mathrm{P}(\mathrm{V},\mu)$ is the set of probability measures π on $X\times X$ such that $\pi(dx)=c(dx)$ and $\pi(X,dy)=\mu(dy)$.One says that μ satisfies the transportation-cost inequality with the cost function $C(x,y)$ if

$$T_c(\mathrm{V},\mu) \le H(\mathrm{V}|\mu), \forall \mathrm{V} \in \mathrm{P}(X) \tag{27}$$

where $H(v|\mu)$ denotes the relative entropy of V with respect to μ and is defined by $H(\mathrm{V}|\mu) = \int \log\left(\frac{d\mathrm{V}}{d\mu}\right)$ if V is absolutely continuous with respect to μ and $H(\mathrm{V}|\mu) = +\infty$ otherwise. Transportation- cost inequalities are known to have good tensorization properties and to yield concentration results independent of the dimension. One will say that μ satisfies the inequality $\mathrm{T}_p(C)$, $\mathrm{P} \in [1,2]$ if it satisfies the transportation cost inequality with the cost function $C(x,y) = \frac{1}{C}\min\left(|x-y|_2^2, |x-y|_2^p\right)$. It is now classical that the inequality $\mathit{T}_p(C)$ implies a concentration inequality similar to (31). When $\rho = 2$, the inequality $TC_2(.)$ is usually denoted by T_2. In [179], M. Talagrand proved that the inequality $TC_2(.)$ is satisfied by Gaussian measures.

In dimension one, an almost complete characterization of transportation-cost inequalities was proposed by the author in [191]. It covers in particular the case of the $TC_p(.)$ inequalities for all $p \in [1,2]$.

In higher dimensions, one only knows that $TC_p(.)$ inequalities and modified logarithmic Sobolev inequalities are related:

(i) For $\rho = 2$, a celebrated result by F. Otto and C. Villani shows that the Logarithmic-Sobolev inequality implies $TC_2(.)$ (see [69]). It was shown by p. Cattiaux and A. Guillin in [204] that the implication is strict: there exist probability measures satisfying $TC_2(.)$ and not the Logarithmic Sobolev inequality. F. Y. Wang provides extensions of Otto and Villani’s result to Riemannian manifolds and path spaces in [72, 74].

(ii) The case $\rho = 1$ is very interesting. S. G. Bobkov, I. Gentil and M. Ledoux have shown in [75] that the inequality $TC_1(.)$ is equivalent to the Poincar´e inequality $\mathrm{SG}(.)$ (see Theorem (6.2.18) for a precise statement).

(iii) For $\mathrm{P} \in [1,2]$, it was shown by I. Gentil, A. Guillin and L. Miclo in [101] that the modified Logarithmic Sobolev inequality $\mathrm{LS}(H_q,.)$ with $\frac{1}{p} + \frac{1}{q} = 1$ implies the transportation-cost inequality $TC_p(.)$.

(iv)The case $\rho > 2$ is much less known. Examples of probability measures satisfying the transportation-cost inequality with a cost function of the form $\left(|x-y|_p^p\right)$ appear in [257] or [33].

Another very efficient functional approach to the concentration of measure phenomenon was proposed by B. Maurey in [17]: the so called(τ) property also called inf-convolution inequality. As inf-convolution inequalities are in fact equivalent to transportation-cost inequalities (see Proposition (6.2.24)

The map w is defined on R but we will also denote byω the map defined on $\mathbf{R}^m$ (for every$m \geq 1$) by $(x_1,...,x_m) \mapsto (w(x_1),...,w(xm))$. The image of a probability measure μ on a space X under a measurable map $T : X \to y$ will be denoted by $T^{\#}\mu$. We recall that it is defined by

$$T^{\#}\mu(B) = \mu\left(T^{-1}(A)\right), \forall A \in y$$

Our paper is organized as follows.

We first recall some well known facts about Poincar´e inequalities. We explains then how to derive general Talagrand's concentration results from the inequalities$\mathrm{SG}(w,C)$ for some $C > 0$, then μ^n concentrates independently of the dimension in the following way: for all $n \geq 1$ and all $A \subset \left(\mathbf{R}^d\right)^n$, one has

$$\mu^n\left(A + B_\omega(h)\right) \geq 1 - e^{-lh}, \forall h \geq 0$$

where $L$is a constant depending only on C and $B_\omega(h)$is the Orlicz ball defined by

$$B_\omega(h) = \left\{(x_1,...,x_n) \in \left(\mathbf{R}^d\right)^n : \sum_{i=1}^{n}\sum_{j1}^{d} \alpha o w\left(\left|\frac{x_{i,j}}{2}\right|\right) \leq h\right\}.$$

(For all $i \leq n, x_{i,j}, 1 \leq j \leq d$ are the coordinates of the vector $x_i \in \mathbf{R}^d$) Proposition (6.2.6) easily implies Proposition (6.2.7) for the special case of the functionsω_p.

We address the problem of finding workable sufficient conditions for the Poincare' inequalities$\mathrm{SG}(w,.)$. To do so, we relate the inequality$\mathrm{SG}(w,.)$ to the classical Poincar´e inequality$\mathrm{SG}(.)$. We show in Proposition (6.2.10) that

$$\mu \text{ verifies} \mathrm{SG}(w,C) \Leftrightarrow w^{\#}\mu \text{ verifies } \mathrm{SG}(C). \tag{28}$$

So, according to (28), to prove that a probability measure μ verifies $\mathrm{SG}(w,.)$, all we have to do is to apply to the measure $w^{\#}\mu$ one of the known criteria for the classical Poincar´e inequality $\mathrm{SG}(.)$. In dimension one, one thus easily derive from the celebrated Muckenhoupt Theorem a necessary and sufficient condition for the inequality $\mathrm{SG}(w,.)$ (see Proposition (6.2.10) Using this criteria, one can give a large collection of examples. Under mild regularity conditions, one proves in Proposition (6.2.13) that a symmetric probability $d\mu(x)=e^{-v(x)}dx$ on $\mathbf{R}$ satisfies the inequality $\mathrm{SG}(w,C)$ for some C if and only if

$$\liminf_{x\to+\infty}\frac{V'(x)}{w'(x)}>0 \tag{29}$$

The same strategy can be applied in dimension d. It is well known that a probability $d\mu(x)=e^{-v(x)}dx$ on $\mathbf{R}$ satisfies the Poincar´e inequality as soon as $\lim_{|x|\to+\infty}\ln f\frac{1}{2}|\nabla V|_2(x^2)-\Delta V(x)>0$ Combined with (28), this criteria yields a sufficient condition for the inequality $\mathrm{SG}(w,.)$ (see Proposition (2.6.12) .We show the equivalence between the Poincare' inequalities for the metrics d_ω and certain transportation-cost inequalities.

Definition (6.2.2) [189]: Let us say that $\mu\in\mathrm{P}(\mathbf{R}^d)$ satisfies the inequality $\mathrm{TC}(w,a)$ if it satisfies the transportation-cost inequality (29) with the cost function $(x,y)\mapsto\alpha(ad_\omega(x,y))$, where $d_\omega(x,y)$ is defined in (22).

In Theorem (6.2.22) which is one of the main results of this section, one proves that μ satisfies the inequality $\mathrm{SG}(w,C)$ for some C if and only if it satisfies the nequality $T\mathrm{C}(w,a)$ for some a.The link between α and C is made precise in Theorem (6.2.22) This theorem is an extension of a result by Bobkov, Gentil and Ledoux concerning the equivalence of the classical Poincare´ inequality and the inequality $T\mathrm{C}_1(.)$(see [256]). This extension is performed using a very simple contraction principle for transportation-cost inequalities. This technique was previously used by [191] to characterize a large class of transportation-cost inequalities on the real line. Since the inequality $T\mathrm{C}(w_p,.)$ is easily shown to be stronger than $T\mathrm{C}_p(.)$, Theorem (6.2.22) offers new sufficient conditions for the transportation-cost inequalities $T\mathrm{C}_p(.)$(see Corollary (6.2.19) .Up to now, Corollary (6.2.19) gives the weakest known sufficient condition for $T\mathrm{C}_p$ inequalities.

We compare the inequalities $\mathrm{SG}(w_p,.)$ to other functional inequalities.

The main result, Theorem (6.2.36) states that under not very restrictive conditions on the function β, the super Poincar´e inequality ($\mathrm{SP}(\beta)$) implies an inequality $\mathrm{SG}(w_p,.)$ where w_β depends only on the function β .Since a lot of functional ine qualities are encoded as super Poincar´e inequalities, this result is extremely general.

As a consequence, one deduces in particular the following relationships.

For $p \in [1,2]$.

$$\mu \text{ verifies } \mathrm{Lo}(p,.) \Leftrightarrow w^{\#}\mu \text{ verifies } \mathrm{SG}$$

Moreover, a counter example of Cattiaux and Guillin shows that the Logarithmic-Sobolev inequality (which corresponds to $\rho = 2$) is strictly stronger than the inequality $\mathrm{SG}(w_2,.)$.

For $p \le 2$ and $1/p + 1/q = 1$, μ verifies $\mathrm{LS}(H_q,.) \Rightarrow \mu$ verifies $\mathrm{SG}(w_p,.)$.

Let us emphasize another interesting fact about Theorem (6.2.36).

We know that super Poincare' inequalities don't tensorize properly. If μ verifies a super Poincar´e inequality, then μ^n will satisfy a super Poincare' inequality with $\beta(s)$ replaced by $\beta(s/n)$. Thus the inequalities deteriorate when the dimension increases. On the other hand, the inequality $\mathrm{SG}(w_\beta,.)$ implied by the super Poincar´e inequality has a good tensorization property and implies concentration independent of the dimension. From this follows that super Poincare' inequalities (almost) always imply dimension free concentration estimates.

Let us recall the two classical structural properties of Poincare' inequalities: tens- orization property and stability under bounded perturbations.

Proposition (6.2.3) [189]: Let μ be a probability on $\mathbf{R}^d$ satisfying the Poincare' inequality $\mathrm{SG}(w,C)$ for some constant $C > 0$.

(i) For all $n > 1$, the probability measure μ^n verifies $\mathrm{SG}(w,C)$ on $(\mathbf{R}^d)^n$

(ii) If $\bar{\mu}$ is a probability measure on $\mathbf{R}^d$ absolutely continuous with respect of μ with a density of the form $d\bar{\mu}(x) = e^{h(x)}d\mu(x)$ with h bounded, then $\bar{\mu}$ verifies the Poincar´e inequality $\mathrm{SG}(w, e^{Osc(h)}C)$ where $Osc(h) = \sup(h) - \inf(h)$.

We will find a proof (in the general case) in [174],[255].

Another way to express the concentration of the product measure μ^n is given in the following corollary which can be easily deduced from the preceding theorem:

Lemma (6.2.4) [189]:(i) For all $x, y \in \mathbf{R}, |w(x) - w(y)| \geq w\left(\frac{|x-y|}{2}\right)$

(ii) The function $x(u) = \min(|u|, u^2)$ is such that $\alpha(xu) \geq \alpha(a)x(u)$, for all $a, u, \geq 0$

Proof: Let us prove the first point. The function $x \to w(x)/x$ is nondecreasing on $\mathbf{R}^+$. It follows that w is super additive on $\mathbf{R}^+$. Indeed, if $0 < x \geq y$ then

$$\begin{aligned} w(x+y) &= w(y(1+x/y)) \geq (1+x/y)w(y) \\ &= w(y) + xw(y)/y \geq w(y) + xw(x)/x = w(y) + w(x) \end{aligned}$$

Let $x \geq y$. If $x \geq y \geq 0$ then using the super additivity of w, one gets

$w(x) = w((x-y)+y) \geq w(x-y) + w(y)$, so $w(x) - w(y) \geq w(x-y) \geq w((x-y)/2)$.

The case $0 \geq x \geq y$ is similar.Now, if $x \geq 0 \geq y$, then

$$w(x) - w(y) = w(x) + w(-y) \geq w(\max(x-y)) \geq w((x-y)/2).$$

Now let us prove the second point. If $0 < a \leq 1, \alpha(au)/\alpha(a) = u^2$, if $u \leq 1/a$ and $\alpha(au)/\alpha(a) = u/a$ if $u \leq 1$ one has $\alpha(au)/\alpha(a) = \alpha(u)$. If $u \in [1, 1/a]$. Then $u^2 \geq u$ and so $\alpha(au)/\alpha(a) \geq \alpha(u)$.If $u \geq 1/a$ then $u/a \geq u$ and so $x(au)/x(a) \geq \alpha(u)$. The case $a \geq 1$ can be handled in a similar way.

Corollary(6.2.5)[274]:(i) For all $x_n, x_{n+1} \in \mathbf{R}, |w(x_n) - w(x_{n+1})| \geq w\left(\frac{|x_n - x_{n+1}|}{2}\right)$

(ii) The function $x_n(u) = \min(|u|, u^2)$ is such that $\alpha(x_{n+1}u) \geq \alpha(a)x_{n+1}(u)$, for all $a, u, \geq 0$

Proof: (i) The function $x_{n+1} \to w(x_{n+1})/x_{n+1}$ is nondecreasing on $\mathbf{R}^+$. Consequently w is super additive on $\mathbf{R}^+$. If $x_{n+1} > 0$ *and* $x_n \leq x_{n+1}$ then

$$w\left(x_{n+1}+x_n\right)=w\left(x_n\left(1+x_{n+1}/x_n\right)\right)\geq\left(1+x_{n+1}/x_n\right)w\left(x_n\right)$$
$$=w\left(x_n\right)+xw\left(x_n\right)/y\ \geq w\left(x_n\right)+x_{n+1}w\left(x_{n+1}\right)/x_{n+1}=w\left(x_n\right)+w\left(x_{n+1}\right)$$

If $x_{n+1}\geq x_n\geq 0$ then using the super additivity of w, we get

$w\left(x_{n+1}\right)=w\left(\left(x_{n+1}-x_n\right)+x_n\right)\geq w\left(x_{n+1}-x_n\right)+w\left(x_n\right)$, so that
$w\left(x_{n+1}\right)-w\left(x_n\right)\geq w\left(x_{n+1}-x_n\right)\geq w\left(\left(x_{n+1}-x_n\right)/2\right)$.

Similar for $x_n\leq x_{n+1}\leq 0$.But; now,if $x_{n+1}\geq 0, x_n\leq 0$,then

$$w\left(x_{n+1}\right)-w\left(x_n\right)=w\left(x_{n+1}\right)+w\left(-x_n\right)\geq w\left(\max\left(x_{n+1}-x_n\right)\right)\geq w\left(\left(x_{n+1}-x_n\right)/2\right).$$

(ii) For the prove of the last part see the proof of Lamma (6.2.4).

Proposition (6.2.6) [189]: Suppose that μ satisfies $\mathrm{SG}(w,C)$ on $\mathbf{R}^d$ for some $C>0$. Then for all $n>1$ and all $A\subset\left(\mathbf{R}^d\right)^n$ with $\mu^n\left(A\right)\geq\frac{1}{2}$ one has

$$\mu^n\left(A+B_\omega\left(h\right)\right)\geq 1-e^{-lh},\forall h\geq 0$$

Where $L=\alpha\left(\frac{1}{\sqrt{Ck}}\right)/\left(16d\right)$ and $B_\omega\left(h\right)$ is defined by

$$B_w\left(h\right)=\left\{\left(x_1,\ldots x_n\right)\in\left(\mathbf{R}^d\right)^n:\sum_{i=1}^{n}\sum_{j=1}^{d}\alpha\circ w\left(\frac{\left|x_i,_j\right|}{2}\right)\leq h\right\}\qquad(30)$$

(For all $1\leq i\leq n, x_i, j, 1\leq j\leq d$ are the coordinates of the vector $x_i\in\mathbf{R}^d$)

Proof : First, $dw\left(u,\mathrm{V}\right)\geq\frac{1}{\sqrt{d}}\sum_{i=1}^{n}\left|w\left(u_i\right)-w\left(\mathrm{V}_i\right)\right|$ For all $u,\mathrm{V}\in\mathbf{R}^d$.

Now,

$$\alpha\left(d_w\left(u,\mathrm{V}\right)\right)\geq\alpha\left(\sum_{i=1}^{n}\frac{1}{\sqrt{d}}\left|w\left(u_i\right)-w\left(\mathrm{V}_i\right)\right|\right)\geq^{(i)}\sum_{i=1}^{d}\alpha\left(\frac{1}{\sqrt{d}}\left|w\left(u_i\right)-w\left(\mathrm{V}_i\right)\right|\right)$$
$$\geq^{(ii)}\sum\alpha\left(\frac{1}{\sqrt{d}}w\left(\frac{\left|u_i-\mathrm{V}_i\right|}{2}\right)\right)\geq^{(iii)}\frac{1}{d}\sum_{i=1}^{d}\alpha o w\left(\frac{\left|u_i-\mathrm{V}_i\right|}{2}\right)$$

where (i) comes from the super additivity of the function x, (ii) from Lemma (6.2.4) (i) and (iii) from Lemma (6.2.4) (ii).

Consequently, for all $x\in\left(\mathbf{R}^d\right)^n$ and $A\subset\left(\mathbf{R}^d\right)^n$.

$$\inf_{a\in A}\sum_{i=1}^{n}\alpha(d_w(x_{i},a_i)) \geq \frac{1}{d}\inf_{a\in A}\sum_{i=1}^{n}\sum_{j=1}^{d} \alpha o w\left(\frac{|x_{i,j}-a_{i,j}|}{2}\right)$$

Applying (40) yields immediately the desired result.

Proposition (6.2.7) [189]: Let $w_p(x)=\max(x,x^p)$ on $\mathbf{R}^+$ with $w_p(-x)=-w_p(x)$ for all $x\in\mathbf{R}$. Suppose that μ satisfies the inequality SG (w_p,C) on $\mathbf{R}^d$ for some $C>0$. If $p\in[1.2]$, then for all $n\geq 1$ and all $A\subset(\mathbf{R}^d)^n$ with $\mu^n(A)\geq\frac{1}{2}$.

$$\mu^n\left(A+2\sqrt{h}B_2+2h\ {}^1\!/\!_p\, B_p\right)\geq 1-e^{-lh},\forall h\geq 0 \tag{31}$$

If $p\geq 2$, then for all $n\geq 1$ and all $A\subset(\mathbf{R}^d)^n$ with $\mu^n(A)\geq\frac{1}{2}$,

$$\mu^n\left(A+2\sqrt{h}B_2\cap 2h^{\frac{1}{p}}B_p\right)\geq 1-e^{-lh},\forall h\geq 0 \tag{32}$$

where L is a constant depending only on C and the dimension d ; one can take $L=\alpha\left(\frac{1}{\sqrt{Ck}}\right)/(16d)$, where $k=\sqrt{18e^{\sqrt{5}}}$

Proof : Suppose $p\in[1,2]$; in view of Proposition (6.2.7) it is enough to prove that

$$\sum_{k=1}^{nd}\alpha o w p(u_k)\leq h\Rightarrow u=(u_1,\ldots,u_{nd})\in\sqrt{h}B_2+h^{1/p}B_p$$

Let $s=(s_1,\ldots,s_{nd})$ and $t=(t_1,\ldots,t_{nd})$ be defined by $s_k=u_k$ if $u_k\in[-1,1]$ and $s_k=0$ if $|u_k|>1$ and $t=u-s$. Then,

$$\sum_{k=1}^{nd}\alpha o w p(u_k)=|s|_2^2+|t|_p^p\leq h$$

So, $|s|_2\leq\sqrt{h}$ and $|t|_p\leq h^{1/p}$. Since $u=s+t$, one concludes that $u\in\sqrt{h}B_2+h^{1/p}B_p$.

Now, if $p\geq 2$, then $\forall x\geq 0,\alpha o w p(x)=\max(x^2,x^p)$. This observation together with Proposition (6.2.6) easily implies the result.

Let us conclude this section with a remark concerning centering. If μ is a probability measure on $\mathbf{R}^d$ and $\mathbf{z} \in \mathbf{R}^d$, let us denote by μ_z the translate of μ defined by:

$$\mu_z(A) = \mu(A+z) \tag{33}$$

for all measurable set A.

The following corollary is immediate.

Corollary (6.2.8) [189]: Suppose that there is some $\mathbf{z} \in \mathbf{R}^d$ such that μ_z verifies the inequality $\mathrm{SG}(w,C)$ for some $C>0$, then for all $n \geq 1$ and all $A \subset (A^d)^n$ with $\mu^n(A) \geq \frac{1}{2}$ one has

$$\mu^n\left(A + B_\omega(h)\right) \geq 1 - e^{-lh}, \forall h \geq 0,$$

Where $L = x\left(\frac{1}{\sqrt{C_k}}\right)/16d$ and $Bw(h)$ is defined by (34).

Definition (6.2.9) [189]: One will say that μ verifies the centered Poincare' inequality $\mathrm{SG}(w,C)$ if $\mu_{\int x d\mu}$ verifies the inequality $\mathrm{SG}(w,C)$.

In order to obtain sufficient conditions for the inequalities $\mathrm{SG}(w,.)$, one relates them to (weighted) forms of the classical Poincare' inequality, which is quite well known.

Proposition (6.2.10) [189]: Let μ be a probability measure on $\mathbf{R}^d$ and C a positive number. The following properties are equivalent.

(i) The probability measure μ verifies $\mathrm{SG}(w,C)$.

(ii) The probability measure $w^{\#}\mu$ verifies $\mathrm{SG}(C)$.

(iii) The probability measure μ satisfies the following weighted Poincare' inequality:

$$Var_\mu(f) \leq C \int \sum_{i=1}^{d} \frac{1}{w'(x_i)^2} \left(\frac{\partial f}{\partial x_i}(x)\right)^2 d\mu(x) \tag{34}$$

for all $f: \mathbf{R}^d \to \mathbf{R}$ such that fow^{-1} is of class C^1.

Proof: Let us denote $|\nabla f|_\omega\left(resp.|\nabla f|_2\right)$ the length of the gradient computed with respect to the metric $d_w(.,.)$ (see (17)). If $f:\mathbf{R}^d\to\mathbf{R}$ is locally Lipschitz for the Euclidean metric, then according to Rademacher theorem, we have

$$\limsup_{y\to x}\frac{|f(x)-f(y)|}{|x-y|_2}=\left[\sum_{i=1}^{d}\left(\frac{\partial f}{\partial x_i}\right)^2(x)\right]^{\frac{1}{2}}=|\nabla f|_2(x)$$

For $\mu a.e.x\in\mathbf{R}^d$, and so the length of the gradient equals the norm of the vector $\nabla f\,\mu a.e$.

Locally Lipschitz function for $d_\omega(.,.)$ and $|.|_2$ are related in the following way.

A function $g:\mathbf{R}^d\to\mathbf{R}$ is locally Lipschitz for $d_\omega(.,.)$ if and only if gow^{-1} is locally Lipschitz for $|.|_2$.

[(i))⇒(ii)] Define $\bar{\mu}=w^{\#}\mu$. Let $f:\mathbf{R}^d\to\mathbf{R}$ be locally Lipschitz fo$|.|_2$, then $f\circ w$ is locallyLipschitz for $d_\omega(.,.)$, and

$$Var_\mu(f):Var_\mu(f\circ w)\le\int\left|\nabla(f\circ w)\right|_\omega^2 d\mu\overset{(*)}{=}\int|\nabla f|_2^2\circ w d\mu=\int|\nabla f|_2^2 d\bar{\mu}$$

where (*) follows from the easy to check identity: $\left|\nabla(f\circ w)\right|_\omega=|\nabla f|_2\circ w$

$\left[(ii)\Rightarrow(i)\right]$The proof is the same.

$\left[(ii)\Rightarrow(iii)\right]$Take $f:\mathbf{R}^d\to\mathbf{R}$ such that $f\circ w^{-1}$ is of class C^1. Then

$$Var_\mu(f)=Var_\mu\left(fow^{-1}\right)\le\int\left|\nabla\left(fow^{-1}\right)\right|_2^2 owd\mu=\int\sum_{i=1}^{d}\frac{1}{w'(x_i)^2}\left(\frac{\partial f}{\partial x_i}\right)^2 d\mu(x)$$

$\left[(iii)\Rightarrow(ii)\right]$Apply the weighted Poincare'inequality to the function $f\circ w$ with f of class C^1.

In the following proposition, a necessary and sufficient condition is given for $\mathrm{SG}(w,.)$ inequalities.

Proposition (6.2.11) [189]: A probability measure μ on $\mathbf{R}$ absolutely continuous with density $h>0$ satisfies the inequality $\mathrm{SG}(w,C)$ for some $C>0$ if and only if

$$D_w^- = \sup_{x \le m} \mu(-\infty, x) \int_x^m \frac{w'(u)^2}{h(u)} du < +\infty$$

and
$$D_w^+ = \sup_{x \ge m} \mu[x, +\infty) \int_m^x \frac{w'(u)^2}{h(u)} du \qquad (35)$$

where m denotes the median of μ. Moreover the optimal constant C in (16) denoted by Copt verifies

$$\max(D_w^-, D_w^+) \le C_{opt} \le 4 \max(D_w^-, D_w^+)$$

This proposition follows at once from the celebrated Muckenhoupt condition for the classical Poincare' inequality (see [218]).

Proof: According to Muckenhoupt condition, a probability measure $dv = hdx$ having a positive continuous density with respect to Lebesgue measure, satisfies the classical Poincare' inequality if and only if

$$D^- = \sup_{x \le m} \mathrm{V}(-\infty, x) \int_x^m \frac{1}{h(u)} du < +\infty \, and D^+ = \sup_{x \ge m} \mathrm{V}[x, +\infty) \int_m^x \frac{1}{h(u)} du < \infty$$

and the optimal constant C_{opt} verifies $\max(D^-, D^+) \le Copt \le 4 \max(D^-, D^+)$ .Now, according to Proposition (6.2.10) μ satisfies $\mathrm{SG}(w, C)$ if and only if $\tilde{\mu} = w^{\#}\mu$ satisfies $\mathrm{SG}(C)$. The density of $\tilde{\mu}$ is $\tilde{h} = \frac{h \circ w^{-1}}{w' \circ w^{-1}}$. Plugging $\tilde{h}$ in Muckenhoupt conditions gives immediately the announced result.

The following result completes the picture giving a large class of examples:

Proposition (6.2.12) [189]: Let μ be an absolutely continuous probability measure on $\mathbf{R}$ with density $du(x) = e^{-v(x)} dx$. Assume that the potential V is of class C^1 and that w verifies the following regularity condition:

$$\frac{w''(x)}{w'^2(x)} \underset{x \to +\infty}{\rightarrow} 0$$

If V is such that

$$\liminf_{x \to \pm\infty} \frac{sgn(x) V'(x)}{w'(x)} > 0 \qquad (36)$$

Then the probability measure μ verifies the Poincare' inequality $\mathrm{SG}(w, C)$ for some $\mathrm{C} > 0$.

Proof: Let $\tilde{\mu} = w^{\#}\mu$ and let V be the symmetric exponential probability measure on $\mathbf{R}$, that is the probability measure with density $d\mathrm{V}(x) = \frac{1}{2}e^{-|x|}dx$. It is well known that it verifies the following Poincare' inequality:

$$Var(g) \le 4\int g'^2(x)\, d\mathrm{V}(x) \qquad (37)$$

for all smooth g [255]. Let $T:\mathbf{R}\to\mathbf{R}$ be the map defined by $T(x) = F_{\tilde{\mu}}^{-1} o F_v(x)$ with $F_v(x) = v(-\infty, x]$ and $F_{\tilde{\mu}}(x) = \tilde{\mu}(-\infty, x]$. It is well known that T is increasing and transports V on $\tilde{\mu}$ which means that $T^{\#}\mathrm{V} = \tilde{\mu}$. Let us apply inequality (37) to a function $g = f \circ T$. It yields immediately:

$$Var_{\tilde{\mu}}(f) \le 4\int f^{12}\left(To'T^{-1}\right)^2 d\tilde{\mu} \le 4\left(\sup_{x\in\mathbf{R}} T'(x)\right)^2 \int f^{12} d\bar{\mu}$$

As a conclusion, if the map T is L Lipschitz then $\tilde{\mu}$ verifies Poincare' inequality $\mathrm{SG}(4L^2)$. The probability $\tilde{\mu}$ has density $d\tilde{\mu}(x) = e^{-\tilde{v}(x)}dx$ with

$$\tilde{V}(x) = \left(w^{-1}(x)\right) + \log w' \circ w^{-1}(x) \ .$$

$\tilde{\mu}$ has density $d\tilde{\mu}(w(x)) = e^{-V(w(x))}dw(x), \tilde{V}(w(x)) = V(x) + \log w'(x)$. The derivative of $\tilde{V}(w(x))\, w.r.t \;\; x$

$$\frac{d\tilde{V}(w(x))}{dw(x)} \cdot \frac{dw(x)}{dx} = \frac{dV}{dx} + \frac{1}{w'(x)}\frac{dw'(x)}{dx}$$

So that

$$\frac{d\tilde{V}}{d(w(x))} = \frac{dV}{dx} \Big/ \frac{dw(x)}{dx} + \frac{1}{w'}\frac{dw(x)}{dx} \Big/ \frac{dw(x)}{dx}$$

It is proved in [191] that a sufficient condition for T to be Lipschitz is that $\liminf_{x\to\pm\infty} \mathrm{sgn}(x)\tilde{V}(x) > 0$. But $\tilde{V}'(w(x)) = \frac{v'(x)}{w'(x)} + \frac{w''(x)}{w^{12}(x)}$ and by assumption $\frac{w''(x)}{w^{12}(x)} \to o$ when x goes to ∞.

Thus

$$\liminf_{x\to\pm\infty} \operatorname{sgn}(x)\tilde{V}(x) = \liminf_{x\to\pm\infty} \frac{\operatorname{sgn}(x)\tilde{V}(x)}{w'(x)}$$

which completes the proof.

Proposition (6.2.13) [189]: Let μ be a probability measure on $\mathbf{R}^d$ absolutely continuous with respect to the Lebesgue measure, with $d\mu(x) = e^{-v(x)}dx$ with V a function of class C^3. Suppose that w is of class C^3 on $\mathbf{R}$ and such that $w'(o) >$ and

$$\left| \frac{w^{(3)}}{(w')^3}(x) \right| \le M \text{ , } \forall x \in \mathbf{R}$$

for some $M \ge 0$. If there is some constant $u > 0$ such that

$$\liminf_{|x|\to\pm\infty} \frac{1}{u^2} \sum_{i=1}^{d} \left[\frac{1}{10} \left(\frac{\partial V}{\partial x_i} \right)^2 \left(\frac{y}{u} \right) - \frac{\partial^2 V}{\partial x_i^2} \left(\frac{y}{u} \right) \right] \frac{1}{w'(x_i)^2} > dM$$

Then the probability measure μ satisfies $\mathrm{SG}(\tilde{w}, C)$ for some C, where $\tilde{w}(x) = w(ux)$, for all $c \in \mathbf{R}$

Proof: It is well known that a probability $dv(x) = e^{-w(x)}dx$ on $\mathbf{R}^d$ satisfies the classical Poincare' inequality if w verifies the following condition:

$$\liminf_{|x|\to\pm\infty} \frac{1}{2} |\nabla w|_2^2(x) - \Delta w(x) > 0 \qquad (38)$$

This condition is rather classical; a nice elementary proof can be found in [26].

Suppose that μ is an absolutely continuous probability measure on $\mathbf{R}^d$ with density $d\mu(x) = e^{-v(x)}dx$ with V of class C^2. Then $\bar{\mu} = w^{\#}\mu$ has density $d\tilde{\mu}(x) = e^{-\tilde{v}(x)}dx$, with

$$\tilde{V}(x) = V\left(w^{-1}(x)\right) + \sum_{i=1}^{d} \log w' \circ w^{-1}(x_i), \forall x \in \mathbf{R}^d$$

According to Proposition (6.2.10) to show that μ satisfies the inequality $\mathrm{SG}(w, C)$ for some $C > 0$ it is enough to show that $\tilde{\mu}$ satisfies the inequality $\mathrm{SG}(C)$ and a sufficient condition for this is that $\tilde{V}$ fulfills condition (38).

Elementary computations yield[274]

$$\nabla\tilde{V}\left(w\left(x\right)\right)=\frac{\partial\tilde{V}}{\partial x_i}\left(w\left(x\right)\right)=\frac{1}{w'\left(x_i\right)}\frac{\partial V}{\partial x_i}\left(x\right)+\frac{w''\left(x_i\right)}{w'^2\left(x_i\right)},$$

$$\Delta\tilde{V}\left(w\left(x\right)\right)=\frac{\partial^2\tilde{V}}{\partial x_i}\left(w\left(x\right)\right)=\frac{w''\left(x_i\right)}{w'^3\left(x_i\right)}\frac{\partial V}{\partial x_i}\left(x\right)+\frac{1}{w'^2\left(x_i\right)}\frac{\partial^2 V}{\partial x_i^2}\left(x\right)+\frac{w^{(3)}\left(x_i\right)}{w'^3\left(x_i\right)}-2\frac{w''^2\left(x_i\right)}{w'^4\left(x_i\right)}$$

Let $1\left(x\right)=\frac{1}{2}\left|\nabla\tilde{V}\right|_2^2\left(w\left(x\right)\right)-\Delta\tilde{V}\left(w\left(x\right)\right),$ we have:

$$1\left(x\right)=\sum_{i=1}^{d}\frac{1}{2}\frac{1}{w'^2\left(x_i\right)}\left(\frac{\partial V_{(x)}}{\partial x_i}\right)^2+\frac{w''\left(x_i\right)}{w'^3\left(x_i\right)}\frac{\partial V}{\partial x_i}+\frac{1}{2}\frac{w''^2\left(x_i\right)}{w'^4\left(x_i\right)}+\frac{w''\left(x_i\right)}{w'^3\left(x_i\right)}\left(\frac{\partial V_{(x)}}{\partial x_i}\right)-\frac{1}{w'^2\left(x_i\right)}\frac{\partial^2 V}{\partial x_i^2}\left(x\right)$$

$$-\frac{w^{(3)}\left(x_i\right)}{w'^3\left(x_i\right)}-2\frac{w''^2\left(x_i\right)}{w'^4\left(x_i\right)}$$

$$=\sum_{i=1}^{d}\frac{1}{w'^2\left(x_i\right)}\left[\frac{1}{2}\left(\frac{\partial V}{\partial x_i}\right)^2\left(x\right)\right]+2\sum_{i=1}^{d}\frac{w''\left(x_i\right)}{w'^3\left(x_i\right)}\frac{\partial V}{\partial x_i}\left(x\right)+\frac{5}{2}\sum_{i=1}^{d}\frac{w''^2\left(x_i\right)}{w'^4\left(x_i\right)}-\sum_{i=1}^{d}\frac{w^{(3)}\left(x_i\right)}{w'^3\left(x_i\right)}$$

Using the inequality $u\mathrm{V}\geq\frac{5}{2}u^2-\frac{1}{5}\mathrm{V}^2,$ one has

$$2\sum_{i=1}^{d}\frac{w''\left(x_i\right)}{w'^3\left(x_i\right)}\frac{\partial V}{\partial x_i}\left(x\right)=2\sum_{i=1}^{d}\left(\frac{w''\left(x_i\right)}{w'^2\left(x_i\right)}\right)\cdot\left(\frac{1}{w'\left(x_i\right)}\frac{\partial V}{\partial x_i}\left(x\right)\right)\geq-\frac{5}{2}\sum_{i=1}^{d}\frac{w''^2\left(x_i\right)}{w'^4\left(x_i\right)}-\frac{2}{5}\sum_{i=1}^{d}\frac{1}{w'^2\left(x_i\right)}\frac{\partial V}{\partial x_i}\left(x\right)$$

and so

$$I\left(x\right)\geq\sum_{i=1}^{d}\frac{1}{w'^2\left(x_i\right)}\left[\frac{1}{2}\left(\frac{\partial V}{\partial x_i}\right)^2\left(x\right)-\frac{\partial^2 V}{\partial x_i^2}\left(x\right)\right]-\frac{5}{2}\sum_{i=1}^{d}\frac{w''^2\left(x_i\right)}{w'^4\left(x_i\right)}-\frac{2}{5}\sum_{i=1}^{d}\frac{1}{w'^2\left(x_i\right)}\left(\frac{\partial V}{\partial x_i}\right)\left(x\right)$$

$$+\sum_{i=1}^{d}\frac{w''^2\left(x_i\right)}{w'^4\left(x_i\right)}-\sum_{i=1}^{d}\frac{w^{(3)}\left(x_i\right)}{w'^3\left(x_i\right)}$$

$$I\left(x\right)\geq\sum_{i=1}^{d}\frac{1}{w'\left(x_i\right)}\left[\frac{1}{10}\left(\frac{\partial V}{\partial x_i}\right)^2\left(x\right)-\frac{\partial^2 V}{\partial x_i^2}\left(x\right)\right]-\sum_{i=1}^{d}\frac{w^{(3)}\left(x_i\right)}{w'^3\left(x_i\right)}$$

Since, $\liminf_{|x|\to+\infty} I(x) = \liminf_{|y|\to+\infty} \frac{1}{2}\left|\nabla \tilde{V}\right|_2^2 (y) - \Delta \tilde{V}(y)$ and

$\sum_{i-1}^{n} \frac{w^3(x_i)}{w'^3(x_i)} \le dM$ .We concludes that $\tilde{V}$ satisfies (38) as soon as

$$\liminf_{|x|\to\pm\infty} \sum_{i=1}^{d} \frac{1}{w'^2(x_i)} \left[\frac{1}{10}\left(\frac{\partial V}{\partial x_i}\right)^2 (x) - \frac{\partial^2 V}{\partial x_i^2}(x) \right] > dM$$

Applying this latter condition to the probability measure $\mu u = (uId)^{\#} \mu$ (where Id is the identity function) which has density $d\mu_u(x) = \frac{1}{u^d} e^{-v(x/u)} dx$ .gives

$$\liminf_{|x|\to+\infty} \frac{1}{u^2} \sum_{i=1}^{d} \frac{1}{w'^2(x_i)} \left[\frac{1}{10}\left(\frac{\partial V}{\partial x_i}\right)^2 \left(\frac{y}{u}\right) - \frac{\partial^2 V}{\partial x_i^2}\left(\frac{y}{u}\right) \right] > dM$$

Where $x = \frac{y}{u}$

Let us recall the notation relative to this family of inequalities. A probability measure μ satisfies the transportation-cost inequality with the cost function $C(x,y)$ on $\mathbf{R}^d$ if for all probability measure V on $\mathbf{R}^d$, the following holds:

$$\inf_{\pi\in P(v,\mu)} \int C(x,y)\, d\pi(x,y) \le H(v|\mu) \qquad (39)$$

where $\mathrm{P}(\mathrm{V},\mu)$ is the set of all probability measures on $\mathbf{R}^d \times \mathbf{R}^d$ such that $\pi(dx\times\mathbf{R}^d) = v(dx)\ and\ \pi(\mathbf{R}^d\times dy) = \mu(dy)\ and\ H(v|\mu)$ is the relative entropy of V with respect to μ.

Wne writes for short that μ satisfies the inequality $T\mathrm{C}(w,a)$ if there is some $a>0$ such that

$$\inf_{\pi\in P(\mathrm{V},\mu)} \int \alpha\left(ad_w(x,y)\right) d\pi(x,y) \le H(\mathrm{V}|\mu), \forall \mathrm{V}$$

with $\alpha(u) = \min(|u|, u^2)$ and $d_w(.,.)$ the distance defined by (22). The purpose of this section is to show that the inequalities $\mathrm{SG}(w,.)$ are equivalent to transportation-cost inequalities $T\mathrm{C}(w,.)$.Transportation-cost inequalities of the form $T\mathrm{C}(w,.)$ are quite unusual. Let us define another family of transportation-cost inequalities appearing often in (see [179,256,101]).Let $p \geq 1$, one says that μ verifies the inequality $TC_p(C)$ if when $p \in [1,2]$

$$\inf_{\pi \in P(V,\mu)} \int \min\left(|x-y|_2^2 |x-y|_2^p\right) d\pi(x,y) \leq CH\left(V|\mu\right), \forall V$$

when $p \in [2,\infty]$

$$\inf_{\pi \in P(V,\mu)} \int \max\left(|x-y|_2^2 |x-y|_2^p\right) d\pi(x,y) \leq CH\left(V|\mu\right), \forall V$$

As we will see, the inequality $TC_p(.)$ is slightly weaker than the inequality $TC_p\left(w_p,.\right)$ (see the proof of Corollary (6.2.19) So in this case, our characterization of inequalities $TC_p(.)$ in terms of Poincar´e inequalities brings new information and criteria for the study of the $TC_p(.)$.

Like Poincare' inequalities, transportation-cost inequalities enjoy a tensorization property and are related to Talagrand's concentration inequalities.

Proposition (6.2.14) [189]: (Tensorization). Suppose that a probability measure μ on a space X satisfies the transportation-cost inequality (39) with the cost function $c(x,y)$,then μ^n satisfies the transportation-cost inequality on X^n with the cost function $c^{\oplus n}(x,y) = \sum_{i=1}^{n} c(x_i, y_i)$. In other words,

$$\inf_{\pi \in P(V,\mu)} \int \sum_{i-1}^{n} c(x_i, y_i)\, d\pi \leq H\left(V|\mu\right), \forall V \in P\left(X^n\right)$$

where $P\left(V,\mu^n\right)$ is the set of probability measures on $X^n \times X^n$ such that $\pi\left(dx, X^n\right) = V(dx)$ and $\pi\left(X^n, dy,\right) = \mu^n(dy)$

This result goes back to the first works of K. Marton on the subject (see [138, 139]). A proof can be found in [190].

Let us explain how to derive concentration inequalities from the inequality $TC(w,a)$.

We will need the following lemma:

Lemma (6.2.15)[189] : The function $x(u) = \min\left(u|u^2\right)$ is such that
$x(x+y) \leq 2\left(x(x)+x(y)\right)$, for all $x, y \geq 0$.

Proof: If $x+y\le 1$ then $\alpha(x+y)=(x+y)^2\le 2(x^2+y^2)=2(\alpha|x)+\alpha(y)$

Now, suppose that $x+y\ge 1$.

If $x\le 1$ and $y\le 1$, then

$$\alpha(x+y)=x+y\le (x+y)^2\le 2(x^2+y^2)=2(\alpha(x)+\alpha(y)).$$

If $x\le 1$ and $y\ge 1$, then

$$x\le y\Rightarrow x-2x^2\le y\Rightarrow x+y\le 2(x^2+y^2)\Rightarrow \alpha(x+y)\le 2(\alpha(x)+\alpha(y))$$

If $x\ge 1$ and $y\ge 1$ then $\alpha(x+y)=x+y=\alpha(x)+\alpha(y)$

$$\le 2(\alpha(x)+\alpha(y))$$

Proposition (6.2.16) [189]: If μ satisfies the transportation-cost inequality $TC(w,a)$, then for all $n\ge 1$ and all $A\subset(\mathbf{R}^d)^n$.

$$\mu^n(A+B_\omega(h))\ge 1-\frac{1}{\mu^n(A)}e^{-h\alpha(a/\sqrt{d})/2},\forall h\ge 0$$

where $B_\omega(h)$ is defined as in Proposition (6.2.7)

Proof : If μ satisfies $TC(w,a)$ on $\mathbf{R}^d$ then according to Proposition (6.2.14) μ^n satisfies the transportation-cost inequality on $\mathbf{R}^d$ with the cost function c defined by

$$c:((x_1,\ldots,x_n),(y_1,\ldots y_n))\in(\mathbf{R}^d)^n\times(\mathbf{R}^d)^n\mapsto\sum\alpha(ad_\omega(x_i,y_i)).$$

Using the triangle inequality for the metric $d_\omega(.,.)$ and Lemma (6.2.15) one has

$$c(x,z)\le 2c(x,y)+2c(y,z),\forall x,y,z\in(\mathbf{R}^d)^n.$$

Now, let ν_1 and ν_2 be two probability measures on $(\mathbf{R}^d)^n$

Take $\pi_1\in P(\nu_1,\mu^n)$ and $\pi_2\in P(\mu^n,\nu_2)$ then one can construct three random variables X,Y,Z such that $L(X,Y,)=\pi_1$ and $L(Y,Z)=\pi_2$ (see [23]).Then, one has

$$T_c(\nu_1,\nu_2) \le \mathrm{E}\left[c(X,Z)\right] \le 2\mathrm{E}\left[c(X,Y)\right] + 2\mathrm{E}\left[c(Y,Z)\right]$$
$$= 2\int c(x,y)d\pi_1(x,y) + 2\int c(y,z)d\pi_2(y,z)$$

Optimizing on π_1 and π_2 gives

$$T_c(\nu_1,\nu_2) \le 2T_c(\nu_1,\mu^n) + 2T_c(\nu_2,\mu^n)$$

Consequently, μ^n satisfies the following symmetrized transportation-cost inequality: for all ν_1,ν_2 probability measures on $(\mathbf{R}^d)^n$.

$$T_c(\nu_1,\nu_2) \le 2H(\nu_1|\mu^n) + 2H(\nu_2|\mu^n)$$

Take $dv_1 = \mathrm{I}_A d\mu^n(A)$ and $dv_2 = \mathrm{I}_A d\mu^n\left|\mu^n(\tilde{A})\right.$ for some $A,\tilde{A} \subset (\mathbf{R}^d)^n$ then

$$\inf_{x\in A, y\in\tilde{A}} c(x,y) \le T_c(\nu_1,\nu_2) \le 2H_c(\nu_1|\mu^n) + 2H(\nu_2|\mu^n)$$
$$= 2\log\left(1/\mu^n(A)\right) + 2\log\left(1/\mu^n(\tilde{A})\right)$$

Letting $c(A,\tilde{A}) = \inf_{x\in A, y\in\tilde{A}} c(x,y)$ we get

$$\mu^n(A)\mu^n(\tilde{A}) \le e^{-c(A,\tilde{A})/2}$$

Defining e $\tilde{A} = \left\{y : \inf c(x,y) > \alpha\left(a/\sqrt{d}\right)h\right\}$ one gets

$$\mu^n(\tilde{A}) \le \frac{1}{\mu^n(A)} e^{-\alpha\left(a/\sqrt{d}\right)h/2}.$$

To obtain the announced inequality it is thus enough to compare $A + B_\omega(h)$ and $\tilde{A}$. Take $x = (x_1,\dots,x_n) \in (\mathbf{R}^d)^n$ and $y = (y_1,\dots,y_n) \in (\mathbf{R}^d)^n$ then for all $i \in 1,\dots,n$, one has

$$\alpha\left(a\mathrm{W}(x_i,y_i)\right) \overset{(a)}{\ge} \alpha\left(\frac{a}{\sqrt{d}}\sum_{i=1}^{d} /w(x_{i,j}) - w(y_{i,j})\right) \overset{(b)}{\ge} \sum_{i=1}^{d}\alpha\left(\frac{a}{\sqrt{d}} /w(x_{i,j}) - w(y_{i,j})\right)$$
$$\overset{(c)}{\ge} \sum_{i=1}^{d}\alpha\left(\frac{a}{\sqrt{d}} \Big/ w\left(\frac{x_{i,j}-y_{i,j}}{2}\right)\right) \overset{(d)}{\ge} \alpha\left(\frac{a}{\sqrt{d}}\right)\sum_{i=1}^{d}\alpha\circ w\left(\frac{x_{i,j}-y_{i,j}}{2}\right)$$

where (a) follows from the comparison between the norms $|.|_2$ and in $|.|_1$ $\mathbf{R}^d$, (b) from the super additivity of x, (c) from Lemma (6.2.4) (i) and (d) from Lemma (6.2.4) (ii).

Consequently, if $y \notin A + B_\omega(h)$ then $\inf \sum_{i=1}^{n}\sum_{j=1}^{d} \alpha \circ w\left(\frac{x_{i,j} - y_{i,j}}{2}\right) \geq h$

and so y belongs to $\tilde{A}$. From this follows that

$$\mu^n\left(A + B_\omega(h)\right) \geq \mu^n\left(\tilde{A}^c\right) \geq 1 - \frac{1}{\mu^n(A)} e^{-\alpha\left(a/\sqrt{d}\right)h/2},$$

which completes the proof.

Corollary (6.2.17)[274]: The function $x_{n+1}(u) = \min\left(u \middle| u^2\right)$ is such that $x_{n+1}\left(x_{n+1} + x_n\right) \leq 2\left(x_{n+1}\left(x_{n+1}\right) + x_{n+1}\left(x_n\right)\right)$, for all $x_{n+1}, x_n \geq 0$.

Proof: If $x_{n+1} + x_n \leq 1$ then

$\alpha\left(x_{n+1} + y\right) = \left(x_{n+1} + x_n\right)^2 \leq 2\left(x_{n+1}{}^2 + x_n{}^2\right) = 2\left(\alpha\left(x_{n+1}\right) + \alpha\left(x_n\right)\right)$

Now, if $x_{n+1} + x_n \geq 1$. If $x_{n+1} \leq 1$ and $x_n \leq 1$, then

$\alpha\left(x_{n+1} + x_n\right) = x_{n+1} + x_n \leq \left(x_{n+1} + x_n\right)^2 \leq 2\left(x_{n+1}{}^2 + x_n{}^2\right) = 2\left(\alpha\left(x_{n+1}\right) + \alpha\left(x_n\right)\right)$

If $x_{n+1} \leq 1$ and $x_n \geq 1$, then

$x_{n+1} \leq x_n$ *implies that* $x_{n+1} - 2x_{n+1}{}^2 \leq x_n$

implies that $x_{n+1} + x_n \leq 2\left(x_{n+1}{}^2 + x_n{}^2\right)$ *implies that* $\alpha\left(x_{n+1} + x_n\right) \leq 2\left(\alpha\left(x_{n+1}\right) + \alpha\left(x_{n+1}\right)\right)$

If $x_{n+1} \geq 1$ and $x_n \geq 1$,

then $\alpha\left(x_{n+1} + x_n\right) = x_{n+1} + x_n = \alpha\left(x_{n+1}\right) + \alpha\left(x_n\right) \leq 2\left(\alpha\left(x_{n+1}\right) + \alpha\left(x_n\right)\right)$

The proof of Theorem (6.2.18) relies on two ingredients. The first one is the following result by Bobkov, Gentil and Ledoux ([256]).

Theorem (6.2.18) [189]: (Bobkov, Gentil, Ledoux). If an absolutely continuous probability measure μ satisfies the inequality $\mathrm{SG}(c)$ on $\mathbf{R}^d$ then it satisfies the transportation-cost inequality for the cost function $(x, y) \mapsto x_2\left(|x - y|_2\right)$ for all $s > \frac{2}{\sqrt{c}}$, where

$$\alpha_s(t) = \begin{cases} \dfrac{t^2}{4L(s)} & if\ |t| \le 2L(s) swithL(s) = \dfrac{C}{2}\left(\dfrac{2+\sqrt{C}s}{2=\sqrt{C}s}\right)^2 e^{s\sqrt{5s}} \\ s|t| - L(s)s^2 & otherwise \end{cases}$$

In particular, if one takes $s = \frac{1}{\sqrt{C}}$, then it is easy to check that $\alpha_s(t) \ge \alpha\left(\frac{1}{\sqrt{Ck}}\right)$, where $\alpha(u) = \min(|u|, u^2)$ and $K = \sqrt{18e^{\sqrt{5}}}$. Thus if μ satisfies $\mathrm{SG}(c)$ it satisfies the transportation-cost inequality with the cost function $(x.y) \mapsto \alpha\left(\frac{|x-y|_2}{\sqrt{Ck}}\right)$. In other words, with the definition of the transportation-cost inequality $\mathrm{SG}(\omega, a)$, the preceding result can be restated as follows.

Corollary (6.2.19) [189]: If μ is an absolutely continuous probability measure on $\mathbf{R}^d$ satisfying the classical Poincar´e inequality $\mathrm{SG}(c)$ for some $C > 0$, then it satisfies the transportation-cost inequality $\mathrm{TC}\left(Id, \frac{1}{\sqrt{Ck}}\right)$. (where $Id : \mathbf{R} \to \mathbf{R} : x \mapsto x$ is the identity function.)

The converse is also true:

Proposition (6.2.20) [1 89]: If μ satisfies $TC(Id, a)$, for some $a > 0$, then μ satisfies the inequality $\mathrm{SG}\left(\frac{1}{2a^2}\right)$.

The proof of Proposition (6.2.20) is classical and can be found in various places (see [256] or [17]).The second argument is a very simple contraction principle:

Proposition (6.2.21) [189]: Let μ be a probability measure on a metric space X ; if μ satisfies the transportation-cost inequality with the cost function $c : XxX \to \mathbf{R}^+$ and if $T : X \to Y$ is a measurable bijection then, $T^{\#}\mu$ satisfies the transportation-cost inequality with the cost function

$$(x, y) \mapsto c\left(T^{-1}(x), T^{-1}(y)\right).$$

This contraction principle goes back to Maurey's work on infimum convolution inequalities (see [17]). A proof can also be found in [191], where this simple property was intensively used to derive necessary and sufficient conditions for transportation-cost inequalities on the real line.

Now let us apply the contraction principle together with Theorem (6.2.18) to prove that Poincare' inequalities $\mathrm{SG}(w,.)$ and transportation-cost inequalities $TC(w,.)$ are qualitatively equivalent.

Theorem (6.2.22) [189]: Let μ be a probability measure on $\mathbf{R}^d$ absolutely continuous with respect to Lebesgue measure. Then μ satisfies the Poincare' inequality $\mathrm{SG}(w,C)$ for some $\mathrm{C}>0$ if and only if it satisfies the transportation-cost inequality $TC(w,a)$ for some $\mathrm{a}>0$. More precisely,

(i) If μ satisfies $\mathrm{SG}(w,C)$ then it satisfies $TC\left(w,\frac{1}{\sqrt{Ck}}\right)$, with $K=\sqrt{18e^{\sqrt{5}}}$.

(ii) If μ satisfies the inequality $TC(w,a)$, then μ satisfies the inequality

$$\mathrm{SG}\left(w,\frac{1}{2a^2}\right).$$

Corollary (6.2.23) [189]: If an absolutely continuous probability measure μ verifies the inequality $\mathrm{SG}(w,C)$ on $\mathbf{R}^d$, for some C and $p\geq 1$, then

(i) if $\mathrm{P}\in[1,2]$ it satisfies the transportation-cost inequality

$$\inf_{\pi\in p(\mathrm{V},\mu)}\int \min\left(|x-y|_2^2,|x-y|_2^p\right)d\pi(x,y)\leq \frac{4}{\alpha\left(\frac{1}{\sqrt{Cdk}}\right)}H(\mathrm{V}/\mu),\forall v$$

(ii) if $\mathrm{P}\geq 2$ it satisfies the transportation-cost inequality

$$\inf_{\pi\in p(\mathrm{V},\mu)}\int \max\left(|x-y|_2^2,|x-y|_2^p\right)d\pi(x,y)\leq \frac{4}{\alpha\left(\frac{1}{\sqrt{Cdk}}\right)}H(\mathrm{V}/\mu),\forall \mathrm{V}$$

Proof: Let $c_p(x,y)=\sum_{i=1}^d \alpha\circ w_p\left(\frac{x_i-y_i}{2}\right)$. During the proof of Proposition (6.2.16) [189]: we have shown that $\alpha\left(a/\sqrt{d}\right)C_p(x,y)\leq\alpha\left(adw_p(x,y)\right)$

So, if μ satisfies the inequality $\mathrm{SG}(w_p,a)$, it satisfies the transportation-cost inequality with the cost function $\alpha\left(a/\sqrt{d}\right)c_p(x,y)$

For $p\in[1,2]$, the function $\alpha o\omega_p\left(\sqrt{.}\right)$ is concave, so

$$c_p(x,y) \geq adw_p\left(|x,y|_2/2\right) \geq \frac{1}{4}\min\left(|x,y|_2^2, |x,y|_2^p\right), \text{for } p \geq 2.$$

$$c_p(x,y) \geq \max\left(\frac{1}{4}|x,y|_2^2, \frac{1}{2_p}|x,y|_p^p\right) \geq 1/2_p \max\left(|x,y|_2^2, |x,y|_p^p\right)$$

The result follows from Theorem (6.2.22)

Let us say that a probability measure μ on a metric space X satisfies the inf-convolution inequality with the cost function $c: X \times X \mapsto \mathbf{R}^+$, if the following holds for all measurable non negative functions $f: X \to \mathbf{R}^+$

$$\int e^{Q_c f} d\mu . \int e^{-f} d\mu \geq 1 \tag{40}$$

where the inf-convolution operator Q_c is defined by

$$Q_c f(x) = \inf_{y \in X}\left\{f(y) + C(x,y)\right\} \tag{41}$$

One will say that a probability measure μ on $\mathbf{R}^d$ satisfies the inf-convolution inequality $\text{IC}(w,a)$ if it satisfies the inf-convolution inequality (40) with the cost function $c(x,y) = \alpha(adw(x,y))$. The inequalities $TC(w,.)$ and $\text{IC}(w,.)$ are qualitatively equivalent, as shown by the following proposition:

Proposition (6.2.24) [189]: If μ verifies the inequality $\text{IC}(w,a)$ then it verifies the inequality $TC(w,a)$.

Conversely, if μ verifies the inequality $TC(w,a)$ then it verifies the inf-convolution inequality with the cost function $2x\left(\frac{a}{2}dw(x,y)\right)$; in particular, it satisfies the inequality $\text{IC}\left(w,\frac{a}{2}\right)$.

Proof: Let $Q^a f(x) = \inf_{y \in X}\left\{f(y) + \alpha\left(ad_\omega(x,y)\right)\right\}$. If μ verifies the inequality $\text{IC}(w,a)$ then, applying Jensen inequality, it holds:

$$\int e^{Q^a f} d\mu \leq e^{\int f d\mu} \tag{42}$$

for all bounded measurable $f: \mathbf{R}^d \to \mathbf{R}$. According to [193], this latter inequality is equivalent to the transportation-cost inequality $\text{TC}(w,a)$.

Conversely, suppose that μ verifies the transportation-cost inequality $\mathrm{TC}(w,a)$. According to [193], the inequality (42) holds. Applying (42) to Q^{af} instead of f, we get

$$\int e^{Q^a(Q^a f)} d\mu . e^{-\int Q^a f d\mu} \leq 1$$

and applying again (42) with $-Q^{af}$ instead of f, we get

$$\int e^{Q^a(-Q^a(f))} d\mu . e^{\int Q^a f d\mu} \leq 1$$

Multiplying these two inequalities yields to

$$\int e^{(Q^\alpha f)} d\mu . \int e^{(-Q^\alpha(f))} d\mu . \leq 1$$

Now, for all $x, y \in \mathbf{R}^d$,we have: $-f(x)+Q^a f(y) \leq \alpha(adw(x,y))$, and consequently, $-f(x) \leq Q^a(-Q^a(f))(x)$. Plugging this into the last inequality gives

$$\int e^{Q^a(Q^{af})} d\mu . \int e^{-f} d\mu \leq 1$$

An easy computation gives:

$$Q^a(Q^a f)(x) = \inf\left\{f(y) + 2\alpha\left(\frac{a}{2} dw(x,y)\right)\right\}$$

This completes the proof.

The following corollary is an immediate consequence of Theorem (6.2.22)

Corollary (6.2.25)[189]: Let μ be a probability measure on $\mathbf{R}^d$ absolutely continuous with respect to Lebesgue measure with a positive density. Then μ satisfies the Poincare' inequality $\mathrm{SG}(w,C)$ for some $C>0$ if and only if it satisfies the inequality $\mathrm{IC}(w,a)$ for some a > 0.More precisely,

(i) If μ satisfies $\mathrm{SG}(w,C)$ then it satisfies $\mathrm{IC}\left(w, \frac{1}{2\sqrt{Ck}}\right)$ with $K = \sqrt{18e^{\sqrt{5}}}$.

(ii)If μ satisfies the inequality $\mathrm{IC}(w,a)$, then μ satisfies the inequality $\mathrm{SG}\left(w, \frac{1}{2a^2}\right)$.

We show that the Poincar´e inequalities $SG(w,.)$ are weaker than super Poincare' inequalities.

Let us recall that μ verifies the super Poincar´e inequality $SP(\beta)$ if for every locally Lipschitz f on $\mathbf{R}^d$, one has

$$\int f^2 d\mu \le \beta(s)\int |\nabla f|_2^2 d\mu + s\left(\int |f| d\mu\right)^2 . \forall s \ge 1 \tag{43}$$

where $\beta:[1,+\infty]\to\mathbf{R}^+$ is nonincreasing.

Indeed, taking $s=1$ in (43) and applying it to $(f-m)_+$, where m denotes the median of the function f, gives:

$$\int (f-m)^2 d\mu \le \beta(1)\int_{f\ge m} |\nabla|_2^2 d\mu + \left(\int_{f\ge m} (f-m) d\mu\right)^2$$

$$\le \beta(1)\int_{f\ge m} |\nabla f|_2^2 d\mu + \frac{1}{2}\int_{f\ge m} (f-m)^2 d\mu$$

Thus, $\int_{f\ge m} (f-m)^2 d\mu \le \beta(1)\int_{f\ge m} |\nabla f|_2^2 d\mu$. Doing the same with $(f-m)_-$

yields , $\int_{f\ge m} (f-m)^2 d\mu \le 2\beta(1)\int_{f\ge m} |\nabla f|_2^2 d\mu$. Adding these inequalities gives

$\int (f-m)^2 d\mu \le 2\beta(1)\int |\nabla f|_2^2 d\mu$. Since $Var_\mu(f) \le \int (f-m)^2 d\mu$, this concludes the proof.

As noted by F.Y. Wang in [75], super Poincare' inequalities imply concentration results. This is recalled in the following proposition.

Proposition (6.2.26) [189]: Suppose that μ verifies (43) with a continuous decreasing function β such that $\beta(s\to 0)$ when s goes to $+\infty$ and define $a=\frac{1}{\sqrt{2\beta(1)}}$, then for all 1-Lipschitz function f on $\mathbf{R}^d$ such that $\int f d\mu = 0$, we have:

$$\int e^{\lambda f} d\mu \le \exp\left(\lambda \int_0^\lambda \phi\left(t \vee e^{i\theta} a\right) dt\right) \forall \lambda \ge 0$$

where the function ϕ is defined by

$$\phi(t)=\frac{1}{t^2}\log\left(2\beta^{-1}\left(\frac{1}{2t^2}\right)\right), \forall t>0$$

As a consequence, defining for all $\lambda \geq 0, \Lambda_\beta(\lambda)=\lambda\int_o^\lambda \phi((t\Lambda a)dt)$

and for all $t\geq 0, \Lambda_\beta^*(t)=\sup_{\lambda\geq 0}\{\lambda t-\Lambda_\beta(\lambda)\}$, we have

$$\mu(f \geq t)e^{-\Lambda_\beta^{*-1}}, \forall t \geq 0$$

Moreover, the inverse function of $\wedge_\beta^*$ can be expressed as follows

$$\Lambda_\beta^{*-1}(t)=\int_0^t \psi(u)\,du, \forall t\geq 0$$

Where $\psi:(0,+\infty]\to \mathbf{R}^+$ is defined by:

$$\psi(t)=\begin{cases}\sqrt{2\log(2)\beta(1)} & if\ t\leq \log(2)\\ \sqrt{2\beta\left(\frac{e^t}{2}\right)} & if\ t\leq \log(2)\end{cases}$$

The observation concerning the inverse of Λ_β^* seems to be new and will be very useful in the sequel. The proof below is simpler than the one proposed by Wang in [75].

Proof: Let f be a 1-Lipschitz function with $\int fd\mu=0$ define $Z(\lambda)=\int e^{2f}d\mu$ and $\Lambda(\lambda)=\log Z(\lambda)$. Applying (45) to the function $e^{\lambda f}$ yields:

$$Z(2\lambda)\leq \lambda^2\beta(s)Z(2\lambda)+sZ(\lambda)^2$$

So, if $s>\beta^{-1}(1/\lambda^2)$, we easily get

$$\Lambda(2\lambda)\leq \log\left(\frac{s}{1-\lambda^2\beta(s)}\right)+2\Lambda(\lambda)$$

Since the function Λ is convex, one has $\Lambda(2\lambda)\geq\Lambda(\lambda)+\lambda\Lambda'(\lambda)$, and so

$$\left[\frac{\Lambda(\lambda)}{\lambda}\right]'=\frac{\lambda\Lambda'(\lambda)-\Lambda(\lambda)}{\lambda^2}\leq\frac{1}{\lambda^2}\log\left(\frac{s}{1-\lambda^2\beta(s)}\right) \qquad (45)$$

If $\lambda < 1/\sqrt{2\beta(1)} = a$, then taking $s=1$ in (46) yields

$$\left[\frac{\Lambda(\lambda)}{\lambda}\right]' \leq -\frac{1}{\lambda^2}\log\left(1-\lambda^2\beta(1)\right) \leq 2\log(2)\beta(1) = \phi(a)$$

If $f\lambda \geq \frac{1}{\sqrt{2\beta(1)}} = a$, then taking $s = \beta^{-1}\left(\frac{1}{2\lambda^2}\right)$ in (46) gives

$$\left[\frac{\Lambda(\lambda)}{\lambda}\right]' \leq \phi(\lambda)$$

So, for all $\lambda > 0, \left[\frac{\Lambda(\lambda)}{\lambda}\right]' \leq \phi(\lambda \vee a)$; ince $\Lambda(\lambda)/\lambda \underset{\lambda\to 0}{\to} 0$ one gets the result.

The inequality $\mu(f \geq t) \leq e^{-\Lambda^*_\beta(t)}$ follows at once from the preceding using routine arguments. Now, let us prove the claim concerning the inverse of Λ^*_β. It is easy to check that

$$\int_0^\lambda \phi(u \vee a)\,du = \int_{1/\lambda}^{+\infty} \frac{1}{u^2}\phi\left(\frac{1}{u} \vee a\right)du = \int_{1/\lambda}^{+\infty} \psi^{-1}(u)\,du = -\int_0^{\psi^{-1}(1/\lambda)} \mathrm{V}\psi'(\mathrm{V})\,d\mathrm{V}$$

Now integrating by part yields

$$\int_0^{\psi^{-1}(1/\lambda)} \mathrm{V}\psi'(\mathrm{V})\,d\mathrm{V} = \frac{\psi^{-1}(1/\lambda)}{\lambda} - \int_0^{\psi^{-1}(1/\lambda)} \psi(u)\,du$$

Let $h(\lambda) = \lambda\int_0^t \psi(u)\,du - \Lambda_\beta(\lambda)$, then

$$h(\lambda) = \lambda\int_{\psi^{-1}(1/\lambda)}^{t} \psi(u)\,du + \psi^{-1}(1/\lambda) = \lambda\int_{\psi^{-1}(1/\lambda)}^{t} \left(\psi(u) - 1/\lambda\right)du + t$$

Observing that ψ is decreasing and $\lambda \mapsto \psi^{-1}(1/\lambda)$ is increasing, it is easy to check that the integral term above is always non positive and vanishes when $\lambda = 1/\psi(t)$. We concludes that

$$\sup_{\lambda\geq 0} h(\lambda) = \Lambda^*_\beta\left(\int_0^t \psi(u)\,du\right) = t,$$

which concludes the proof.

Lemma (6.2.27) [189]: Suppose that $\beta:[1,+\infty)\to\mathbf{R}^+$ is continuous decreasing function such that $s\mapsto s\beta(s)$ is nondecreasing on $[1,+\infty)$ and define $w_p:\mathbf{R}^+\to\mathbf{R}^+$ as follows:

$$w_\beta^{-1}(t)=4\int_0^t\sqrt{\beta\left(e^u\right)du},\quad \forall t\geq 0 \tag{45}$$

Then we have

$$\alpha o w_\beta(t)\leq\Lambda_\beta^*(t)\leq\alpha o w_\beta(5t),\quad \forall t\geq 0 \tag{46}$$

Where $\partial(t)=\min\left(t^2,t\right)$ for all $t\geq 0$.

Proof: Let us prove the lower bound in (46). According to Proposition (6.2.26) this inequality is equivalent to the following one

$$\Lambda_\beta^*(\alpha(t))=\int_0^{t^2\Lambda t}\psi(u)\,du\leq 4\int_0^t\sqrt{\beta\left(e^u\right)du},\quad \forall t\geq 0 \tag{47}$$

where the function ψ is defined in Proposition (6.2.26).In fact a slightly better inequality holds true:

$$\Lambda_\beta^*(\alpha(t))=\int_0^{t^2\Lambda t}\psi(u)\,du\leq 2\sqrt{2}\int_0^t\sqrt{\beta\left(e^u/2\right)}du\quad \forall t\geq 0 \tag{48}$$

with the convention $\beta(s)=\beta(1)$, when $s\in[0,1]$. Since the function $s\beta(s)$ is nondecreasing on $[1,+\infty)$ it is easy to check that $\beta\left(e^u/2\right)\leq 2\beta\left(e^u\right)$, and so (48) implies (47). To prove (48), let us distinguish the following cases.

If $t\leq\log(2)$, then

$$\Lambda_\beta^{*-1}(\alpha(t))=\Lambda_\beta^{*-1}\left(t^2\right)=2\sqrt{2\log(2)\beta(1)}t\leq 2\sqrt{2\beta(1)}t=2\sqrt{2}\int_0^t\sqrt{\beta\left(e^u/2\right)}du$$

If $\log(2)\leq t$, then

$$\Lambda_{\beta}^{*-1}\left(\alpha(t)\right) \le \Lambda_{\beta}^{*-1}(t) = 2\sqrt{2\beta(1)}\log(2) + \int_{l\,og(2)}^{t} \sqrt{2\beta\left(e^{u}/2\right)}du$$

$$\le 2\int_{0}^{l\,og(2)} \sqrt{2\beta\left(e^{u}/2\right)}du + \int_{l\,og(2)}^{t} \sqrt{2\beta\left(e^{u}/2\right)}du \le \int_{0}^{t} \sqrt{\beta\left(e^{u}/2\right)}du$$

The proof of the upper bound in (46) is similar.

Examples (6.2.28) [189]: Let $P \ge 1$, and define $\beta_p(s) = \log(e+s)^{2(1/p-1)}$ (which verifies the conditions $s\beta(s)$ increasing according to Lemma (6.2.38) Then, we can show that

$$w_p\left(\frac{t}{4p\left(2^{1/p}-1\right)}\right) \ge w_{\beta_p}(t) \ge w_p\left(\frac{t}{4p}\right), \forall t \ge 0 \qquad (49)$$

where $w_p(u) = u \vee u^p$ for all $u \ge 0$. In particular, if μ verifies inequality $\mathrm{SP}\left(C\beta_p\right)$ for some $C > 0$, then we have

$$\mu\left(|x|_2 \ge t + \int |x|_2\, d\mu\right) \le e^{-\alpha o w_p\left(t/\left(4\sqrt{C}p\right)\right)}, \forall t \ge 0$$

and this implies that $\int e^{\varepsilon|x|_2^p} d\mu < +\infty$ for some $\varepsilon > 0$, Since the probability measure $dv_p(x) = \frac{1}{Z_p} e^{-|x|_p} dx$ verifies $\mathrm{SP}\left(C\beta_p\right)$, for some $C > 0$, one concludes that the function w_{β_p} gives the right order of concentration. We think that more generally the function w_{β} is of the right order.

Now we can state the following result:

Let us recall the definition of a capacity-measure inequality [262]and [59,60]

Definition (6.2.29) [189]: Let μ be a probability measure on $\mathbf{R}^d$. Let $A \subset \Omega$ be Borel sets. One defines

$$Cap_{\mu}(A,\Omega) = \inf\left\{\int |\nabla f|_2^2\, d\mu : \mathrm{I}_A \le \mathrm{I}_{\Omega}\right\}$$

The capacity of a set A with $\mu(A) \le 1/2$ is defined by

$$Cap_\mu(A) = \inf\{Cap_\mu(A,\Omega) : A \subset \Omega \, and \, \mu(\Omega) \le 1/2\}$$
$$= \inf\left\{\int |\nabla f|_2^2 \, d\mu : f : \mathbf{R}^d \to [0,1], f_{|A} = 1 and \mu(f = 0) \ge 1/2\right\}$$

One says that μ satisfies a capacity-measure inequality if there is a function $\psi : [0,1] \to \mathbf{R}^+$ such that for all A with $(\mu(A) \le 1/2), \psi(\mu(A)) \le Cap_\mu(A)$

Many functional inequalities admit a transcription in terms of capacity measure. The simplest example is the classical Poincar´e inequality on $\mathbf{R}^d$.

Theorem (6.2.30) [189]: A probability measure μ on $\mathbf{R}^d$ verifies the inequality

$\mathrm{SG}(C)$ for some $C > 0$ if and only if there is some $D > 0$ such that for all $\mathbf{A} \subset \mathbf{R}^d$ with $\mu(A) \le 1/2$, $\mu(A) \le DCap_\mu(A)$.

Moreover, optimal constants verify $D_{opt}/2 \le C_{opt} \le 4D_{opt}$.

Theorem (6.2.31) [189]: (Barthe-Cattiaux-Roberto). Let $\beta : [1,+\infty] \to \mathbf{R}^+$ be a nonincr-easing function such that $s \mapsto s\beta(s)$ is nondecreasing. Suppose that for all $\mathbf{A} \subset \mathbf{R}^d$, with $\mu(A) \le 1/2$,

$$\frac{\mu(A)}{\beta(1/\mu(A))} \le Cap_\mu(A)$$

then μ verifies the super Poincare' inequality $\mathrm{SP}(8\beta)$.

In fact, for our purpose one is only interested in the converse proposition:

Proposition (6.2.32)[189]: Let $\beta : [1,+\infty] \to \mathbf{R}^+$ be a nonincreasing function such that $s \mapsto s\beta(s)$ is nondecreasing. Suppose also that there exists $\lambda \ge 4$ such that

$$\lambda\beta(\lambda s) \ge 4\beta(s), \forall s \ge 1$$

Under the preceding assumption, if μ verifies the super Poincare' inequality $\mathrm{SP}(\beta)$, then for all $\mathbf{A} \subset \mathbf{R}^d$, with $\mu(A) \le 1/2$ we have

$$\frac{\mu(A)}{\beta(1/\mu(A))} \le 4\lambda Cap_\mu(A)$$

Proof: The following proof is a straightforward adaptation of the proof of [59] and we will only sketch it. Let $\mathbf{A} \subset \mathbf{R}^d$ with $\mu(A) \le 1/2$ and $f : \mathbf{R}^d \to [0,1]$ a function which is 1 on A and vanishes with probability more than 1/2. For all $k \in Z$, define $f_k = (f - 2^k)_+ \Lambda 2^k$ and $\Omega_k = \{f \ge 2^k\}$. Applying the super Poincare' inequality (43) to the function f_k one obtains:

$$f_k^2 d\mu \le \beta(s) \int |\nabla f_k|_2^2 d\mu + s\left(\int |f_k| d\mu\right)^2$$
$$\le \beta(s) \int |\nabla f|_2^2 \, d\mu + s\mu(\Omega_k) \int f_k^2 d\mu$$

Taking $s = \dfrac{1}{2_\mu(\Omega_k)} \ge 1$ and noticing that $f_k^2 \ge 2^{2k}$ on Ω_{k+1} gives

$$\mu(\Omega_{k+1}) 2^{2k} \le \int f_k^2 d\mu \le 2\beta\left(\frac{1}{2_{\mu(\Omega_k)}}\right) \int |\nabla f|_2^2 \, d\mu$$

Defining $F(x) = \dfrac{1}{2\beta(x/2)}$ for $x \ge 2$, $a_k = \mu(\Omega_k)$, and $C = \int |\nabla f|^2 \, d\mu$ one gets

$2^{2k} a_{k+1} F(1/ak) \le \lambda C$, as soon as $a_k = 0$. Applying [59], we concludes that $2^{2k} a_k F(1/a_k) \le \lambda C$ as soon as $a_k = 0$. If one takes $K = 0$, one has $A \subset \Omega o$ so $a_o \ge \mu(A)$ and since $s\beta(s)$ is nondecreasing, $a_o F(1/a_o) \ge \mu(A) F(1/\mu(A))$. Consequently,

$$\frac{\mu(A)}{4\beta\left(\dfrac{1}{\mu(A)}\right)} \le \frac{\mu(A)}{2\beta\left(\dfrac{1}{2\mu(A)}\right)} \le \lambda \int |\nabla f|_2^2 \, d\mu$$

Optimizing over f gives the result.

In all what follows, we will adopt the following convention: for $s \le 1$, one defines $\beta(s) = \beta(1)$.

Let

$$\Theta(x) = \frac{x}{4\lambda\beta(1/k)} \quad \text{for all } x > 0. \tag{50}$$

Where λ is defined in (51).

Lemma (6.2.33) [189]: If $\beta:[1,+\infty]\to\mathbf{R}^{+}$ is a nonincreasing function such that

$s\mapsto s\beta(s)$ is nondecreasing then the function Θ defined by (50) is nondecreasing and verifies $\Theta(x+y)\le\Theta(x)+\Theta(y)$ for all $x,y\in\mathbf{R}^{+}$.

Proof: Since $s\beta(s)$ is nondecreasing, it follows that Θ is nondecreasing. Moreover, since β is nonincreasing, it follows that $\Theta(x)/x$ is nonincreasing. Thus, if $x\ge y>0$; we get

$$\Theta(x+y)=\Theta(x(1+y/x))\le(1+y/x)\Theta(x)$$

$$=\Theta(x)/x\le\Theta(x)+\Theta(y).$$

This completes the proof.

The following lemma explains how behave capacity-measure inequalities under push-forward:

Lemma (6.2.34) [189]: Suppose that μ satisfies the capacity-measure inequality

$$\psi(\mu(A))\le DCap_{\mu}(A), \text{with } \mu(A)\le\frac{1}{2}, \forall A$$

Then $\bar{\mu}=w^{\#}\mu$ verifies the inequality

$$\psi(\tilde{\mu}(A))\le D\overline{Cap}_{\tilde{\mu}}(A), \text{with } \tilde{\mu}(A)\le\frac{1}{2}, \forall A$$

Where

$$\overline{Cap_{\tilde{\mu}}}=\inf\left\{\int\sum_{i=1}^{d}\left(w\circ w^{-1}(x_i)\right)^2\left(\frac{\partial f}{\partial x_i}\right)^2(x)\,d\tilde{\mu}: f\mathbf{R}^d\to[0,1], f_{|A}=1 \text{ and } \tilde{\mu}(f0)\ge\frac{1}{2}\right\}.$$

Proof: Let A be such that $\tilde{\mu}(A)\le\frac{1}{2}$, and f be such that $f=1$ on A and $\tilde{\mu}(f=0)\ge\frac{1}{2}$. Define $B=w^{-1}(A)$ and $g=f\circ w$. Then $\mu(B)=\tilde{\mu}(A)\le\frac{1}{2}$. $g\ge1$ on B and $\{g=0\}=w^{-1}\{f=0\}$

and $\mu(f=0)=\tilde{\mu}(f=0)\ge\frac{1}{2}$. Applying the capacity-measure inequality verified by μ to B and g yields

$$\psi(\tilde{\mu}(A)) = \psi(\tilde{\mu}(B)) \le D\int |\nabla_g|_2^2 \, d\mu = D\int \sum_{i=1}^{d} \left(w' \circ w^{-1}(x_i)\right)^2 \left(\frac{\partial f}{\partial x_i}\right)^2 (x) \, d\tilde{\mu}$$

Optimizing over such functions f gives the announced inequality for $\tilde{\mu}$.

The next lemma compares the capacity $\overline{Cap_{\tilde{\mu}}}$ to the usual capacity Cap_{μ}:

Lemma (6.2.35) [189]: Suppose that w is convex and

Let $B_\infty(r) = \{x \in \mathbf{R}^d : \max_{1 \le i \le d} |x_i| \le r\}$, for all $r \ge 0$. If $A \subset B_\infty(r)$ and $\mu(A) \le \frac{1}{2}$, then

$$Cap_{\tilde{\mu}}(A) \le 2\left(w' \circ w^{-1}(r+1)\right)^2 \left[Cap_{\tilde{\mu}}(A) + \tilde{\mu}\left(B_\infty(r)^c\right)\right]$$

Proof: Let

$$Cap_{\tilde{\mu}}^r(A) = \inf\left\{\int |\nabla f|_2^2 \, d\tilde{\mu} : 1_A \le f \le 1_{B_{\infty(r+1)}} \ and \ \tilde{\mu}(f=0) \ge \frac{1}{2}\right\}$$

Using the fact that the function $\omega' o \omega^{-1}$ is nondecreasing on $\mathbf{R}^+$, we clearly have:

$$\overline{Cap}_{\tilde{\mu}}(A) \le \left(w' \circ w^{-1}(r+1)\right)^2 Cap_{\tilde{\mu}}^r(A)$$

Now let $f : \mathbf{R}^d \to [0,1]$ be such that $f_{|A} = 1$ and $\tilde{\mu}(f=0) \ge \frac{1}{2}$. Let $h : \mathbf{R} \to \mathbf{R}^+$ defined

by $h(t) = (r+1-t)_+ \wedge 1$ and consider $\phi : \mathbf{R}^d \to \mathbf{R}^+$ defined by $\phi(x) = h(|x|_\infty)$. It is

not difficult to check that $|\nabla_\phi|_2 \le 1_{B_\infty(r)^c}$. Let $g = f\phi$; we have

$$1_A \le g \le 1_{B_\infty(r+1)}, \tilde{\mu}(g=0) \ge \tilde{\mu}(f=0) \ge \frac{1}{2},$$

and

$$Cap_{\tilde{\mu}}^r(A) \le \int |\nabla_g|_2^2 \, d\tilde{\mu} = \int |\nabla f \, \phi + f \, \nabla \phi|_2^2 \, d\tilde{\mu} \le 2\int |\nabla f|_2^2 \phi^2 d\tilde{\mu} + 2\int f^2 |\nabla \phi|_2^2$$
$$\le 2\int |\nabla f|_2^2 \, d\tilde{\mu} + 2\tilde{\mu}\left(B_\infty(r)^c\right)$$

Optimizing over f yields:

$$Cap_{\tilde{\mu}}^{r}(A) \leq 2Cap_{\tilde{\mu}}(A) + 2\tilde{\mu}\left(B_{\infty}(r)^{c}\right)$$

Theorem (6.2.36) [189]: Let $\beta : [1, +\infty) \to \mathbf{R}^{+}$ be a continuous decreasing function such that $s \mapsto s\beta(s)$ is increasing and such that there is some $\lambda \geq 4$ for which the following holds

$$\lambda\beta(\lambda_{s}) \geq 4\beta(s), \forall s \geq 1 \tag{51}$$

If a probability measure μ on $\mathbf{R}^{d}$ verifies the super Poincar´e inequality $\mathrm{SP}(\beta)$, then there is some a > 0 such that μ verifies $\mathrm{SG}\left(w_{\beta}(./a), 4\lambda^{2}\right)$, where ω_{β} is defined by (50) for $t \geq 0$ and extended to $\mathbf{R}^{-}$ by $w_{\beta}(t) = -w_{\beta}(-t)$, for $t \neq 0$. We can take

$$a = \max\left(\lambda, \Lambda_{\beta}^{*}(m)\right)$$

Where $m = \int |x|_{2}\, d\mu$

Moreover, under the same assumptions, the probability measure μ verifies the centered Poincare' inequality $\mathrm{SG}\left(w_{\beta}(./\bar{a}), 4\lambda^{2}\right)$(see Definition (6.2.29) with

$$\bar{a} = \max\left(\lambda, \Lambda_{\beta}^{*}\left(\sqrt{2\beta(1)d}\right)\right)$$

The constant $\bar{a}$ above depends only on β and enjoys the following invariant property: if β is replaced by $t\beta$ with $t > 0$, then $\bar{a}$ is unchanged.

Finally, under the same assumptions, the probability measure μ verifies the following transportation-cost inequality

$$\inf_{\pi \in P(\mathrm{V}, \mu)} \int \sum_{i=1}^{d} \alpha \circ w_{\beta}\left(\frac{x_{i} - y_{i}}{2\bar{a}}\right) d\pi(x, y) \leq \frac{1}{\alpha\left(\frac{1}{2\lambda_{k}\sqrt{d}}\right)} H(\mathrm{V}\mu)$$

for all probability measure V on $\mathbf{R}^{d}$.

Proof : Define $\tilde{\mu}$ as the image of μ under the map $x \mapsto w_{p}(x/a)$ with $a = \max\left(\lambda, \Lambda_{\beta}^{*}(m)\right)$. We wants to prove tha $\tilde{\mu}$ verifies the classical Poincare' inequality

According to Proposition (6.2.32) the probability measure μ satisfies the capacity-measure inequality

$$\Theta\left(\mu\left(A\right)\right) \le Cap_{\mu}\left(A\right), with\ \tilde{\mu}\left(A\right) \le \frac{1}{2},\ \forall A. \qquad (52)$$

According to Lemma (6.2.35) $\tilde{\mu}$ satisfies the capacity-measure type inequality:

$$\Theta\left(\tilde{\mu}\left(A\right)\right) \le \overline{Cap}_{\mu}\left(A\right), with\ \tilde{\mu}\left(A\right) \le \frac{1}{2},\ \forall A,$$

where $\overline{Cap}_{\tilde{\mu}}$ is defined in the lemma.

Let $B_{\infty}\left(r\right) = \left\{x \in \mathbf{R}^d : \max_{1\le i\le d}\left(\left|x_i\right|\right) \le r\right\}$, for all $r \ge 0$. Let $A \subset \mathbf{R}^d$ with $\tilde{\mu}\left(A\right) \le \frac{1}{2}$; one has

$$\begin{aligned}
\Theta\left(\tilde{\mu}\left(A\right)\right) &\overset{(i)}{\le} \Theta\left(\tilde{\mu}\left(A \cap B_{\infty}\left(r\right)\right)\right) + \Theta\left(\tilde{\mu}\left(B_{\infty}\left(r\right)^c\right)\right) \\
&\overset{(ii)}{\le} \overline{Cap}_{\tilde{\mu}}\left(A \cap B_{\infty}\left(r\right)\right) + \Theta\left(\tilde{\mu}\left(B_{\infty}\left(r\right)^c\right)\right) \\
&\overset{(iii)}{\le} 2/a^2\left(w'_{\beta} \circ w^{-1}\left(r+1\right)\right)^2\left[Cap_{\tilde{\mu}}\left(A \cap B_{\infty}\left(r\right)\right) + \tilde{\mu}\left(B_{\infty}\left(r\right)^c\right)\right] + \Theta\left(\tilde{\mu}\left(B_{\infty}\left(r\right)^c\right)\right) \\
&\overset{(iv)}{\le} 2/a^2\left(w'_{\beta} \circ w^{-1}\left(r+1\right)\right)^2\left[Cap_{\tilde{\mu}}\left(A\right) + \tilde{\mu}\left(B_{\infty}\left(r\right)^c\right)\right] + \Theta\left(\tilde{\mu}\left(B_{\infty}\left(r\right)^c\right)\right) \\
&\overset{(v)}{\le} \frac{e}{8a^2\beta\left(e^r\right)}\left[Cap_{\tilde{\mu}}\left(A\right) + \tilde{\mu}\left(B_{\infty}\left(r\right)^c\right)\right] + \Theta\left(\tilde{\mu}\left(B_{\infty}\left(r\right)^c\right)\right)
\end{aligned}$$

where (i) follows from the sub-additivity and the monotonicity of Θ, (ii) from Lemma(6.2.33) (iii) from Lemma (6.2.34) and the convexity of w_{β}, (iv) from the fact that the function $A \mapsto Cap_{\tilde{\mu}}\left(A\right)$ is nondecreasing and (v) from the equation (45) of w_{β} and the inequality $\beta\left(e^{r+1}\right) \ge 1/e\beta\left(e^r\right)$. Thanks to Lemma (6.2.37) below, one has

$$\tilde{\mu}\left(\beta_{\infty}\left(r\right)^c\right) \le \left(2e\right)e^{-r}.$$

Using the monotonicity and the sub-additivity of Θ, one has

$$\Theta\left(\tilde{\mu}\left(\beta_{\infty}\left(r\right)^c\right)\right) \le \Theta\left(\left(2e\right)e^{-r}\right) \le \left(2e\right)\Theta\left(e^{-r}\right).$$

So, letting $t = \tilde{\mu}\left(A\right)$ and using the definition of Θ, one has:

$$\frac{t}{\beta(1/t)} \le \left(\frac{e\lambda}{2a^2} Cap_{\tilde{\mu}}(A) + \left(\frac{e^2\lambda}{a^2} + 2e\right)e^{-r}\right)\frac{1}{\beta(e^r)}.\forall r \ge 0$$

Since $a \ge \lambda \ge 4$, one has $\frac{e\lambda}{2a^2} \le 1/2$ and $\frac{e^2\lambda}{a^2} + 2e \le 8$ and so letting $b = \frac{t}{\beta(1/t)}$, one gets

$$\sup_{s\ge1}\{b\beta(s) - 8/s\} \le \frac{1}{2} Cap_{\tilde{\mu}}(A)$$

Let $\mathsf{g}(s) = s\beta(s), s \ge 1$; by hypotheses g is increasing and goes to $+\infty$ when $s \to +\infty$.

Taking $s = g^{-1}(16/b)$ (which is well defined) yields

$$b\beta\left(g^{-1}(16/b)\right) \le Cap_{\tilde{\mu}}(A)$$

According to (52), one has $g(\lambda x) \ge 4g(x)$ for all $x \ge 1$; from this follows that

$g^{-1}(4x) \le \lambda g^{-1}(x)$ for all $x \ge \beta(1)$ and by iteration $g^{-1}(16x) \le \lambda^2 g^{-1}(x)$, for all $x \ge \beta(1)$. Consequently,

$$g^{-1}(16/b) \le \lambda^2 g^{-1}(1/b) = \lambda^2 g^{-1}\left(g(1/t)\right) = \lambda^2/t$$

As β is nonincreasing, one concludes that $\beta\left(\lambda^2/t\right) \le \beta\left(g^{-1}(16/b)\right)$.

Since $\beta\left(\lambda^2/t\right) \ge \beta(1/t)\lambda^2$, one gets $t/\lambda^2 \le b\beta\left(g^{-1}(16/b)\right) \le Cap_{\tilde{\mu}}(A)$.

In other word, for all $A \subset \mathbf{R}^d$ with $\tilde{\mu}(A) \le \frac{1}{2}$

$$\tilde{\mu}(A) \le \lambda^2 Cap_{\tilde{\mu}}(A)$$

According to Theorem (6.2.30) one concludes that $\tilde{\mu}$ verifies the classical

Poincare' inequality $\mathrm{SG}(4\lambda^2)$.

Let $\tilde{\mu} = \mu_{\int xd\mu}$. If μ verifies the super Poincare' inequality (43), then so does $\tilde{\mu}$. So all the preceding results apply to $\tilde{\mu}$. In particular, $\tilde{\mu}$ verifies the inequality

$$\mathrm{SG}\left(w_\beta(./a), 4\lambda^2\right)$$

with $a=\max\left(\lambda,\wedge_{\beta}^{*}(\bar{m})\right)$, where $\bar{m}=\int\left|x-\int y d\mu\right|_{2} d\mu(x)$. But,

$$\bar{m}^{2}\leq\sum_{i=1}^{d}\int\left(x_{i}-\int y_{i} d\mu\right)^{2} d\mu(x)\leq 2\beta(1) d,$$

where the first inequality follows from Cauchy-Schwarz inequality and the second from the fact that μ verifies the Poincar´e inequality $\mathrm{SG}(2\beta(1))$ (see Remark (6.2.41). This proves that $\tilde{\mu}$ verifies $\mathrm{SG}\left(w_{\beta}(./a),4\lambda^{2}\right)$ with

$$\bar{a}=\max\left(\lambda,\Lambda_{\beta}^{*}\left(\sqrt{2\beta(1d)}\right)\right).$$

The invariance property of $\bar{a}$ follows immediately from the definition of Λ_{β}^{*} given in Proposition (6.2.20)

Now, according to Theorem (6.2.22) $\tilde{\mu}$ verifies the inequality $TC\left(w_{\beta}(./\bar{a}),\frac{1}{2\lambda k}\right)$

Reasoning as in the proof of Corollary (6.2.19) one sees that this implies that $\tilde{\mu}$ satisfies the transportation-cost inequality with the cost function

$$c(x,y)=\alpha\left(\frac{1}{2\lambda k\sqrt{d}}\right)\sum_{i=1}^{d}\alpha\circ w_{\beta}\left(\frac{|x_{i}-y_{i}|}{2\bar{a}}\right).$$

Since transportation-cost inequalities are translation invariant, this concludes the proof.During the proof of Theorem (6.2.30) we have used the following lemma.

Lemma (6.2.37) [189]: The probability measure $\tilde{\mu}$ which is the image of μ under the $map\, x\mapsto w_{\beta}(x/a)$ with $a=\max\left(\lambda,\Lambda_{\beta}^{*}(m)\right) and\, m=\int|x|_{2} d\mu$ verifies

$$\tilde{\mu}\left(|x|_{\infty}\geq r\right)\leq(2e)e^{-r},\ \forall r\geq 0$$

Proof: According to Lemma (6.2.33) and e.g [192], one has

$$\int e^{\varepsilon\Lambda_{\beta}^{*}\left(|x|_{2}-m\right)} d\mu\leq\frac{1+\varepsilon}{1-\varepsilon},\quad\forall\varepsilon\in[0,1],$$

where Λ_{β}^{*} is defined in Proposition (6.2.20).Using the convexity of Λ_{β}^{*}

and the fact that $\Lambda_{\beta}^{*}\geq\alpha\circ w_{\beta}$ one gets since $a\geq 2$.

$$\exp\left(\alpha \circ w_{\beta}\left(|x|_2 / a\right)\right) \le \exp\left(\Lambda_{\beta}^{*}\left(|x|_2 / a\right)\right) \le \exp\left(\frac{1}{a}\Lambda_{\beta}^{*}|x|_2 - m\right).\exp\left(\frac{1}{a}\Lambda_{\beta}^{*}m\right)$$

Since $\left|w_{\beta}(x/a)\right|_{\infty} = w_{\beta}\left(|x|_{\infty}/a\right) \le w_{\beta}\left(|x|_2 / a\right)$, integrating yields:

$$\int e^{\alpha\left(|x|_{\infty}\right)} d\,\tilde{\mu}(x) \le \frac{1+1/a}{1-1/a}.exp\left(\frac{1}{a}\Lambda_{\beta}^{*}(m)\right) \le 2e$$

which gives the result.

In this section we will draw consequences of Theorem (6.2.36) .We will focuss on the functions $\beta_p(s) = \log(e+s)^{2(1/p-1)}$, but more general results could be stated. First let us show that these functions verify the assumptions of Theorem (6.2.36).

Lemma (6.2.38) [189]: For all $p \ge 1$, the function $\beta_p(s) = \log(e+s)^{2(1/p-1)}$ is such that

$s \mapsto s\beta_p(s)$ is increasing on $[0,+\infty]$. Moreover, for all $p \ge 1$, there is some $\lambda \ge 4$, such that $\lambda\beta_p(\lambda s) \ge 4\beta_p(s)$ for all $s \ge 1$. Let us denote by λ_p the smallest of these $\lambda's$, then the $map\ p \mapsto \lambda_p$ is increasing. Moreover, one always has $\lambda_p \le 205$ for all $p \ge 1$ and for $p \in [1,2]$, one has $\lambda_p \le 20$.

Proof: Let $r = 2(1-1/p)$; then $r \in [0,2]$. The $map\ s \mapsto \log(e+s)^r$ is concave on $[0,+\infty]$. Consequently, the $map\ s \mapsto \left(\log(e+s)^r - 1\right)/s$ decreases on $[0,+\infty]$ and so does

$s \mapsto \log(e+s)^r / s$. In other word $s \mapsto s\beta_p(s)$ is increasing.

Next observe that $\lambda\beta_p(\lambda s) \ge 4\beta_p(s) \Leftrightarrow \lambda\left[\frac{\log(e+s)}{\log(e+\lambda s)}\right]^r \ge 4$. This clearly implies that the $map\ p \mapsto \lambda_p$ map is nondecreasing.

Let $f(s) = \dfrac{\log(e+s)}{\log(e+\lambda s)}$ then

$$f'(s) = \frac{(e+\lambda s)\log(e+\lambda s) - \lambda(e+s)\log(e+s)}{\log(e+\lambda s)^2 (e+s)(e+\lambda s)} = \frac{\phi(\lambda s) - \lambda\phi(s)}{\log(e+\lambda s)^2 (e+s)(e+\lambda s)}$$

with $\phi(s) = (e+s)\log(e+s)$. Then $\dfrac{d}{ds}\dfrac{(\phi(s))}{s} = \dfrac{s - e\log(e+s)}{s^2}$. If $s \ge 6$, then

$\frac{d}{dx}\left(\frac{\phi(s)}{s}\right)\geq 0$ so $s\mapsto\phi(s)/s$ is nondecreasing and this implies that

$\phi(\lambda s)\geq\lambda\phi(s)$ for all $s\geq 6$. As a consequence, $f'(s)\geq 0$ when $s\geq 6$ and the function f is thus nondecreasing on $[6,+\infty]$. Consequently, $f(s)\geq f(6)$ for $s\geq 6$ and

$f(s)\geq\frac{1}{\log(e+6\lambda)}$ for $s\geq 6$. Since $f(6)\geq\frac{1}{\log(e+6\lambda)}$, one has $f(s)\geq\frac{1}{\log(e+6\lambda)}$

for all $s\geq 1$.

From what precedes one concludes it is enough to find $\lambda\geq 4$ such that

$$\frac{\lambda}{\log(e+6\lambda)}\geq 4$$

For $r=2$, one checks that $\lambda=205$ is convenient and for $r=1$, one can take $\lambda=20$. This the proof.

Let us recall that μ satisfies the Lata la-Oleszkiewicz inequality $\mathrm{LO}(p,C)$ if

$$\sup_{a\in(1,2)}\frac{\int f^2d\mu-\left(\int|f|^a\,d\mu\right)^{2/a}}{(2-a)^{2(1-1/p)}}\leq C\int|\nabla f|_2^2\,d\mu,\quad\forall f.$$

(53)

The following result is due to F. Y. Wang (see [73]):

Theorem (6.2.39) [189]: Let $p\in[1,2]$; a probability measure verifies the $\mathrm{LO}(p,C)$ for some $C>0$ if and only if it verifies the super Poincar´e inequality $\mathrm{SP}\left(\tilde{C}\beta_p\right)$.

Corollary (6.2.40) [189]: If μ verifies the inequality $\mathrm{LO}(p,C)$ on $\mathbf{R}^d$, with $p\in[1,2]$ then μ verifies the centered inequality $\mathrm{SG}\left(w_p\left(./\,a_1\sqrt{C}\right),a_2\right)$, where a_1 depends only on the dimension d and a_2 is an absolute constant. One can take $a_1=4\sqrt{6}$ max (5d, 20) and $a_2=(320)^2$.

Remark (6.2.41) [189]: The fact that the dimension d appears in the constant a_2 above is not a problem, thanks to the tensorization property of the (centered) Poincare' inequality.

Proof: According to Theorem (6.2.39) μ verifies $\mathrm{SP}\left(96C\beta_p\right)$ and according to Theorem (6.2.36) μ verifies the centered Poincare' inequality

$$\mathrm{SG}\left(w, 4\lambda_p^2\right) w := w_{96C\beta_p}\left(./\bar{a}\right) = w_p\left(./\left(4\bar{a}\sqrt{6C}\right)\right).$$

According to Lemma (6.2.38) we have $\lambda_p = 20$. Using the inequalities (46) and (49), one sees that $\Lambda_\beta^*\left(\sqrt{2\beta_p(1)d} \le \Lambda_\beta^*\sqrt{2d}\right) \le 5d$, so $\bar{a} \le \max(5d, 20)$. It is easy to check that $w'\beta_p \ge \frac{1}{4p}w'_p \ge \frac{1}{8}w'_p$ According to Proposition (6.2.10) one concludes that μ verifies the centered Poincare' inequality $\mathrm{SG}\left(w_p./4\max(5d,20)\sqrt{6C},(320)^2\right)$.

Let $H:\mathbf{R}\to\mathbf{R}^+$; let us recall that μ verifies the modified Log-Sobolev inequality

$\mathrm{LS}(H,C)$ on $\mathbf{R}^d$, if for all locally Lipschitz positive function f;

$$H\left(\mu\middle|f^2\right) \le C\sum_{i=1}^{d}\int H\left(\frac{\partial_i f}{f}\right) f^2 d\mu$$

Let $p \ge 2$ define q such that $1/p + 1/q = 1$ and $H_q(x) = |x|^q$. The inequality $\mathrm{LSI}\left(H_q,.\right)$

is related to super Poincar´e inequality $\mathrm{SP}\left(\beta_p\right)$ as explained in the following

Proposition (6.2.42) [189]: Let $P \ge 2$ and suppose that μ verifies the inequality

$\mathrm{LSI}\left(H_q,C\right)$ on $\mathbf{R}^d$ with $1/p + 1/q = 1$, then μ verifies the super Poincare' inequality

$\mathrm{SP}\left(C^{2(1-1/p)}k\beta_p\right)$, where k is a constant depending only on the dimension d and p.

Proof: Since the function $x \mapsto x^{q/2}$ is concave, applying Jensen inequality yields:

$$\int H_q\left(\frac{\partial_i f}{f}\right) f^2 d\mu \le \left(\int (\partial_i f)^2 \, d\mu\right)^{q/2} \left(\int f^2 d\mu\right)^{1-q/2}$$

So, using concavity again,

$$H\left(\mu \middle| f^2\right) \le C d^{1-q/2} \left(\int |\nabla f|_2^2 \, d\mu\right)^{q/2} \left(\int f^2 d\mu\right)^{1-q/2}$$

Since, $x^{q/2} = \inf_{s>0}\left\{sx + a_q s^{\frac{a}{q-2}}\right\}$ with $a_q = \left(\frac{2-q}{2}\right)\left(\frac{2}{q}\right)^{\frac{q}{q-2}}$, one concludes that for all s > 0

$$H\left(\mu \middle| f^2\right) \le \tilde{C} s \int |\nabla f|_2^2 \, d\mu + \tilde{C} a_q s^{\frac{q}{q-2}} \int f^2 d\mu$$

letting $\tilde{C} = C d^{1-q/2}$. According to the proof of [75], if a probability measure μ verifies an inequality of the form:

$$H\left(\mu \middle| f^2\right) \le C_1 \int |\nabla f|_2^2 \, d\mu + C_2 \int f^2 d\mu$$

Then it verifies

$$\int f^2 d\mu \le r \int |\nabla f|_2^2 \, d\mu + \left(\frac{rC_2}{2C_1}\right)^2 \exp\left(C_2 + \frac{2C}{r}\right) \left(\int |f| d\mu\right)^2, \quad \forall r > 0$$

From this follows, that

$$\int f^2 d\mu \le r \int |\nabla f|_2^2 d\mu + \left(\frac{a_q}{2} r s^{\frac{2}{q-2}} + 1\right)^2 \exp\left(\tilde{C} a_q s^{\frac{q}{q-2}} + 2\tilde{C} s / r\right) \left(\int |f| \, d\mu\right)^2$$

holdsfor all $s, r > 0$. Choosing $s = r^{\frac{2-q}{2}}$ yields:

$$\int f^2 d\mu \le r \int |\nabla f|_2^2 \, d\mu + \frac{1}{4}\left(a_q + 2\right)^2 \exp\left(\tilde{C} + \left(a_q + 2\right) r^{-q/2}\right) \left(\int |f| d\mu\right)^2, \quad \forall r > 0$$

or equivalently:

$$\int f^2 d\mu \le \tilde{C}^{2/q} b_q^{1/q} \log\left(\frac{4s}{b_q}\right)^{-2/q} \int |\nabla f|_2^2 \, d\mu + s\left(\int |f| d\mu\right)^2, \quad \forall s \ge \frac{1}{4} b_q \ .$$

Where $b_q=\left(a_q+2\right)^2$. According to [252], μ verifies the Poincare' inequality $\mathrm{SG}\left(c_q c^{2/q}\right)$, where $c_q=36.6^{2/q}$. Let $\beta(s)=c_q\Lambda d^{2/q-1}b_q^{1/q}\left(\frac{4s}{b_q}\right)^{-2/q}$, for $s\geq b_q/4$ and $\beta(s)=c_q$ for $s\in\left[1,b_q/4\right]$, then μ verifies the super Poincare' inequality $\mathrm{SP}\left(C^{2/q}\beta\right)$. It is clear that one can find a constant k such that $\beta\leq k\beta_p$. This constant K depends only on d and q.

Reasoning exactly as in Corollary (6.2.43) we prove the following result.

Corollary (6.2.43) [189]: Let $p\geq 2$ and suppose that μ verifies the inequality $\mathrm{LS}\left(H_q,C\right)$ on $\mathbf{R}^d$ with $1/p+1/q=1$, then μ verifies $\mathrm{SG}\left(w_p\left(./ac^{1-1/p}\right),b\right)$, where a and b are constants depending only on d and p.

Theorem (6.2.44) [189]: (Bobkov-Ledoux). If μ satisfies (16), then for every bounded function f on X^n such that $\sum_{i=1}^{n}\left|\nabla_i f\right|^2\leq a^2$ and $\max\left|\nabla_i f\right|\leq b,\mu^n a.e$ (where $\left|\nabla_i f\right|$ denotes the length of the gradient with respect to the i^{th} coordinate) we have

$$\mu^n\left(f\leq\int f\,d\mu^n+t\right)\leq\exp\left(-\min\left(\frac{t^2}{Ck^2a^2},\frac{t}{\sqrt{C}kb}\right)\right),\forall t\geq 0 \tag{54}$$

With $k=\sqrt{18e^{\sqrt{5}}}$

proof: (i) According to [255] (which is the main result of [255]), μ enjoys a modified Logarithmic-Sobolev inequality: for all $0<s<\frac{2}{\sqrt{C}}$ and for all locally Lipschitz $f:X\to\mathbf{R}$ such that $\left|\nabla f\right|\leq s\,\mu\,a.e$. We have

$$H\left(\mu\middle|e^f\right)\leq L(s)\int\left|\nabla f\right|^2 e^f\,d\mu \tag{55}$$

Where $L(s)=\frac{C}{2}\left(\frac{2+\sqrt{C}s}{2-\sqrt{C}s}\right)^2 e^{s\sqrt{5C}}$

(ii) Tensorization. Thanks to the tensorization property of the entropy functional,

$$H\left(\mu^n \left| e^f \right.\right) \le \int \sum_{i=1}^{n} H_\mu\left(e^{f_i}\right) d\mu^n$$

for all $f : X^n \to \mathbf{R}$.

Applying this inequality together with (40) yields

$$H\left(\mu^n \left| e^f \right.\right) \le L(s) \int \sum_{i=1}^{n} \left|\nabla_i f\right|^2 e^f d\mu . \tag{56}$$

For all $0 < s < \frac{2}{\sqrt{C}}$ and $f : X^n \to \mathbf{R}$ such that $\max_{1 \le i \le n} \left|\nabla_i f\right| \le s\mu^n \ a.e.$

(iii) Herbst argument. Thanks to the homogeneity one can suppose that $f = X^n \to \mathbf{R}$ is such that $\max_{1 \le i \le n} \left|\nabla_i f\right| \le 1 (b = 1)$ and $\sum_{i=1}^{n} \left|\nabla_i f\right|^2 \le a^2$. Define $Z(\lambda) = \int e^{\lambda f} d\mu^n$ Then, applying (56) to λf , we easily obtains the following differential inequality

$$\frac{d}{d\lambda}\left(\frac{\log\left(Z(\lambda)\right)}{\lambda}\right) \le L(s) a^2, \quad \forall 0 < \lambda \le s < \frac{2}{\sqrt{C}}$$

and since $\frac{\log\left(Z(\lambda)\right)}{2} \to \int d\mu^n as\ \lambda \to 0$, we get

$$\int e^{\lambda f} d\mu^n \le e^{\lambda 2 L(s) a^2 + \lambda \int f d\mu^n}, \quad \forall 0 < \lambda \le s < \frac{2}{\sqrt{C}}$$

(iv) Tchebychev argument. This latter inequality on the Laplace transform yields via Tchebychev argument:

$$\mu^n\left(f \ge \int f d\mu^n + t\right) \le e^{-h_s(t)}, \quad \forall t \ge 0$$

Where

$$h(t) = \sup_{\lambda \in [0,s]} \left\{\lambda t - L(s) a^2 \lambda^2\right\} = \begin{cases} \dfrac{t^2}{4L(s)a^2} & if\ 0 \le t \le 2L(s)a^2 s^2 \\ st - L(s)a^2 s^2 & if\ t \ge 2L(s)a^2 s^2 \end{cases}$$

Now it easy to see that , $h_s(t) \ge \min\left(\frac{t^2}{4L(s)a^2}, \frac{st}{2}\right)$. For $s = 1/\sqrt{C}$ one obtains after some computations,

$$h_s(t) \geq \min\left(\frac{t^2}{Ck^2a^2}, \frac{t}{\sqrt{C}k}\right) \quad \text{with } k = \sqrt{18e^{\sqrt{5}}}$$

Corollary (6.2.45) [189]: (Bobkov-Ledoux). Let μ be a probability measure on X satisfying the Poincar´e inequality (16) on (X, d) with the constant $C > 0$. There is a constant L depending only on C such that for all subset A of X^n with $\mu^n(A) \geq \frac{1}{2}$,

$$\mu^n\left(A^h\right) \geq 1 - e^{-lh}, \forall h \geq 0 \tag{57}$$

where the set A^h is the enlargement of A defined by

$$A^h = \left\{ y \in X^n : \inf_{x \in A} \sum_{i=1}^{n} \alpha\left(d\left(x_i, y_i\right)\right) \leq h \right\}$$

where $\alpha(u) = \min\left(|u|, u^2\right)$ for all $u \in \mathbf{R}$. One can take $L = \left(\frac{1}{\sqrt{C}k}\right)/16$ where as before $k = \sqrt{18e^{\sqrt{5}}}$.

proof: Take $A \subset X^n$, such that $\mu^n(A) \geq \frac{1}{2}$ and define $F(x) = \inf_{a \in A} \sum_{i=1}^{n} \alpha\left(d\left(x, a_i\right)\right)$, where $\alpha(u) = \min\left(|u|, u^2\right)$. Then for all $r > 0$, the function $f = \min(F, r)$ verifies (see [255]) $\max_{1 \leq i \leq n} |\nabla_i f| \leq 2$ and $\sum_{i=1}^{n} |\nabla_i f|^2 \leq 4r$ .Moreover since $\mu^n(A) \geq \frac{1}{2}$, we have

$\int f d\mu^n = \int f 1_{A^c} d\mu^n \leq r\left(1 - \mu^n(A)\right) \leq r/2$. Consequently, applying (54) to f yields:

$$\mu^n(F \geq r) = \mu^n(f \geq r) \geq \mu^n\left(f \geq \int d\mu^n + r/2\right) \leq e^{-rK(C)}$$

with $K(C) = \frac{1}{16} \min\left(\frac{1}{Ck^2}, \frac{1}{\sqrt{C}k}\right) = \frac{1}{16} \alpha\left(\frac{1}{\sqrt{C}k}\right)$. This concludes the proof of (57).

Corollary (6.2.46)[274]: If μ is a probability measure satisfies (16), then for every bounded function Lipichitzf^{j} on X^{n} such that $\sum_{j=1}^{n}\sum_{i=1}^{n}\left|\nabla_{i}f^{j}\right|^{2}\leq a^{2}$ and $\max\sum_{j=1}^{n}\left|\nabla_{i}f^{j}\right|\leq b$, $\mu^{n}\,a.e$ we have

$$\mu^{n}\left(\sum_{j=1}^{n}f^{j}\leq\int\sum_{j=1}^{n}f^{j}\,d\mu^{n}+t_{j}\right)\leq\prod_{j=1}^{n}\mathrm{e}^{\left(-\min\left(\frac{t_{j}^{2}}{Ck^{2}a^{2}},\frac{t_{j}}{\sqrt{C}kb}\right)\right)},\forall t_{j}\geq 0$$

With $k=\sqrt{18e^{\sqrt{5}}}$

proof: (i) For all $0<s^{j}<\frac{2}{\sqrt{C}}$ and for all locally Lipschitz$f^{j}:X\to\mathbf{R}$ such that $\sum_{j=1}^{n}\left|\nabla f^{j}\right|\leq s^{j}\,\mu\,a.e$. We get

$$\sum_{j=1}^{n}H\left(\mu\middle|e^{f^{j}}\right)\leq L\left(s^{j}\right)\int\sum_{j=1}^{n}\left|\nabla f^{j}\right|^{2}e^{f^{j}}\,d\mu$$

Such that $L\left(s^{j}\right)=\frac{C}{2}\left(\frac{2+\sqrt{C}s^{j}}{2-\sqrt{C}s^{j}}\right)^{2}e^{s^{j}\sqrt{5C}}$

(ii)According to the tensorization property of the entropy functional,

$$\sum_{j=1}^{n}H\left(\mu^{n}\middle|e^{f^{j}}\right)\leq\int\sum_{j=1}^{n}\sum_{i=1}^{n}H_{\mu}\left(e^{f^{j}_{i}}\right)d\mu^{n}$$

for $f^{j}:X^{n}\to\mathbf{R}$,with (55) give

$$\sum_{j=1}^{n}H\left(\mu^{n}\middle|e^{f^{j}}\right)\leq L\left(s^{j}\right)\int\sum_{j=1}^{n}\sum_{i=1}^{n}\left|\nabla_{i}f^{j}\right|^{2}e^{f^{j}}\,d\mu\,.$$

For $0<s^{j}<\frac{2}{\sqrt{C}}$ and $f:X^{n}\to\mathbf{R}$ s.t $\max_{1\leq i\leq n}\sum_{j=1}^{n}\left|\nabla_{i}f^{j}\right|\leq s^{j}\mu^{n}\;a.e$.

(iii) Given $\max_{1\leq i\leq n}\sum_{j=1}^{n}\left|\nabla_{i}f^{j}\right|\leq 1(b=1)$ and $\sum_{j=1}^{n}\sum_{i=1}^{n}\left|\nabla_{i}f^{j}\right|^{2}\leq a^{2}$. Define $Z\left(\lambda^{j}\right)=\int e^{\lambda^{j}f^{j}}d\mu^{n}$. Applying (56) we have

$$\frac{d}{d\lambda^{j}}\left(\frac{\log\left(Z\left(\lambda^{j}\right)\right)}{\lambda^{j}}\right)\leq L\left(s^{j}\right)a^{2},\quad\forall 0<\lambda^{j}\leq s^{j}<\frac{2}{\sqrt{C}}$$

Since $\frac{\log\left(Z\left(\lambda^j\right)\right)}{2} \to \int d\mu^n$ when $\lambda^j \to 0$, we find

$$\int e^{\lambda^j f^j} d\mu^n \le e^{\lambda^j 2L\left(s^j\right)a^2 + \lambda^j \int f^j d\mu^n}, \quad \forall 0 < \lambda^j \le s^j < \frac{2}{\sqrt{C}}$$

(iv) $\mu^n\left(f^j \ge \int f^j d\mu^n + t^j\right) \le e^{-h_{s^j}\left(t^j\right)}, \quad \forall t^j \ge 0$

Where

$$h\left(t^j\right) = \sup_{\lambda^j \in \left[0, s^j\right]} \left\{\lambda^j t^j - L\left(s^j\right)a^2\left(\lambda^j\right)^2\right\} = \begin{cases} \frac{\left(t^j\right)^2}{4L\left(s^j\right)a^2} & \text{if } 0 \le t^j \le 2L\left(s^j\right)a^2\left(s^j\right)^2 \\ s^j t^j - L\left(s^j\right)a^2\left(s^j\right)^2 & \text{if } t^j \ge 2L\left(s^j\right)a^2\left(s^j\right)^2 \end{cases}$$

Then $h_{s^j}\left(t^j\right) \ge \min\left(\frac{\left(t^j\right)^2}{4L\left(s^j\right)a^2}, \frac{s^j t^j}{2}\right)$, where $k^2 = 18e^{\sqrt{5}}$

Section (6.3): Mass transportation of free functional inequalities and free Poincaré inequalities

A distinguished role in the world of functional inequalities is played by the logarithmic Sobolev (Log-Sobolev) inequality and the Talagrand or transportation cost inequality. There is an extensive literature dedicated to these inequalities in the classical setting of Euclidean and Riemannian spaces (see [25, 174, 23, 71])

Given a probability measure V on $\mathbf{R}^d$, the transportation cost inequality $\left(\mathrm{T}(\rho)\right)$ states that for some $\rho > 0$ and any other probability measure μ on $\mathbf{R}^d$,

$$\rho \mathrm{W}_2^2\left(\mu, \mathrm{V}\right) \le H\left(\mu | \mathrm{V}\right)$$

Here $\mathrm{W}_2\left(\mu, \mathrm{V}\right)$ is the Wasserstein distance between μ and V of finite second moment defined by

$$\rho \mathrm{W}_2\left(\mu, \mathrm{V}\right) = \inf_{\pi \in \prod(\mu, \mathrm{V})}\left(\int\int |x-y|^2 \pi\left(dx, du\right)\right)^{\frac{1}{2}}$$

with $\Pi\left(\mu, \mathrm{V}\right)$ denoting the set of probability measures on $\mathbf{R}^d$ with marginals μ and V and

$$H(\mu|\mathrm{V}) = \int \log \frac{d\mu}{d\mathrm{V}} d\mu$$

is the relative entropy of μ with respect to V if $\mu << \mathrm{V}$and $+\infty$ otherwise. The Log-Sobolev inequality $(\mathrm{LSI}(\rho))$ is that for any μ

$$H(\mu|\mathrm{V}) \leq \frac{1}{2\rho} 1(\mu|\mathrm{V})$$

Where

$$I(\mu|\mathrm{V}) = \int \left| \nabla \log \frac{d\mu}{d\mathrm{V}} \right| d\mu$$

is the Fisher information of μ with respect to V which is defined in the case $\mu << \mathrm{V}$ with $\frac{d\mu}{d\mathrm{V}}$ being differentiable. A more subtle inequality is the $(HWI(\rho))$inequality relating entropy, Wasserstein distance W, and Fisher information I

$$H(\mu|\mathrm{V}) \leq \sqrt{1(\mu|v)} \mathrm{W}_2(\mu, \mathrm{V}) - \frac{\rho}{2} \mathrm{W}_2^2(\mu, \mathrm{V})$$

Poincaré's inequalit$(\mathrm{P}(\rho))$y in this classical context is that for any compactly supported and smooth function ψ on $\mathbf{R}^d$,

$$\rho Var_\mu(\psi) \leq \int |\nabla \psi|^2 \mu(dx)$$

Where $Var_\mu(\psi) = \int \psi^2(x)\mu(dx) - \left(\int \psi(x)\mu(dx)\right)^2$is the variance of ψ with respect to μ.

To wit a little bit here, let $V: \mathbf{R} \to \mathbf{R}$ be a nice function with enough growth at infinity and define the probability distribution

$$\mathrm{P}_n(dM) = \frac{1}{Z_n} e^{-nTn(V(M))} dM$$

On the set H_n of complex Hermitian $n \times n$matrices where dM is the Lebesgue measure onH_n. For a matrix M, let $\mu_n(M) = \frac{1}{n} \sum_{k=1}^{n} \delta_{\lambda k}(M)$be the distribution of eigenvalues of M. These are random variables with values in $\mathrm{P}(\mathbf{R})$, the set of probability measures on $\mathbf{R}$ which converge almost surely to a non-random measure μ_V on $\mathbf{R}$. For a measure μ on $\mathbf{R}$, it's the logarithmic energy with external field V is defined by

$$\mathrm{E}(\mu)=\int V(x)\,\mu(dx)-\iint \log|x-y|\mu(dx)\,\mu(dy).$$

The minimizer of $\mathrm{E}(\mu)$ over all probability measures on $\mathbf{R}$ is exactly the measure μ_V . From [78] we learned that the distributions of $\{\mu_n\}_{n\geq 1}$ under P_n satisfy a large deviations principle with scaling n^2 and rate function given by

$$\mathbf{R}(\mu)=\mathrm{E}(\mu)-\mathrm{E}(\mu_{\mathrm{V}}).$$

The example of the quadratic potential $V(x)=x^2$ defining the paradigmatic Gaussian Unitary Ensemble in random matrix theory gives rise to the celebrated semicircular law as equilibrium measure.

Within this random matrix framework, if $V(x)=px^2$ is smooth and convex for some $\rho>0$, then the function $\Phi(M)=Tr_n(V(M))$ is strongly convex $\Phi(M)-n\rho|M|^2$ is convex) on $\mathbf{R}^{n2}=\mathrm{H}_n$. An application of the classical $LSI(n\rho)$ on H_n for large n was used by Biane [200] to prove a Log-Sobolev inequality in the context of one-dimensional free probability which holds

(cf. [66]) in the following form

$$\mathrm{E}(\mu)-\mathrm{E}(\mu_V)\leq\frac{1}{4\rho}1(\mu) \tag{58}$$

for any probability measure μ on $\mathbf{R}$ whose density with respect to the Lebegue measure is in $L^3(\mathbf{R})$, where

$$I(\mu)=\int\left(H\mu(x)-V'(x)\right)^2\mu(dx)$$

with $H_\mu=2\int\frac{1}{x-y}\mu(dx)$ being the Hilbert transform of μ.

More precisely, Biane and Voiculescu used the free Ornstein Uhlenbeck process and the complex Burger equation. Using the large random matrix strategy, Hiai Petz and Ueda [66] reproved and extended the result of Biane and Voiculescu in the following form. If $V(x)-px^2$ is convex for some $\rho>0$, then for every probability measure μ on $\mathbf{R}$,

$$\rho\mathrm{W}_2^2(\mu,\mu_V)\leq\mathrm{E}(\mu)-\mathrm{E}(\mu_V) \tag{59}$$

For example, we cover potentials V on the line such that $V(x)-\rho|x|^p$ is convex for some $\rho>0$ and $\mathrm{p}>1$ as well as a class of bounded perturbations of convex potentials. Using this approach, we present here an HWI free inequality for various cases of potentials. For the case $V(x)-\rho x^2$ convex for some $p \in \mathbf{R}$, this is

$$\mathrm{E}(\mu)-\mathrm{E}(\mu_V) \leq \sqrt{I(\mu)}\mathrm{W}_2(\mu,\mu_V)-\rho \mathrm{W}_2^2(\mu,\mu_V) \tag{60}$$

Also a Brunn–Minkowski inequality receives a direct proof as well.

The second part of this work is devoted to free one-dimensional Poincaré inequalities. Using random matrix approximations and the classical Poincaré inequality, we first give an ansatz to what could be a possible Poincaré inequality in the free probability world. In the case of $V(x)-\rho x^2$ convex for some $\rho > 0$, such that the measure μ_v has support $[-1,1]$, this states as,

$$\int \phi'(x)^2 \mu_V(dx) \geq \frac{\rho}{2\pi^2}\int_{-1}^{1}\int_{-1}^{1}\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2 \frac{1-xy}{\sqrt{1-x^2}\sqrt{1-y^2}}dxdy \tag{61}$$

for any smooth function ϕ on the interval [−1, 1].

There is also a second version of the Poincaré which is discussed in [200] for the case of the semicircular law. This inequality has a natural meaning in the context of free probability as the derivative $\nabla\phi$ of a function from the classical $\mathrm{P}(\rho)$ is replaced by the noncommutative derivative $\frac{\phi(x)-\phi(y)}{x-y}$, and thus our second version takes the form

$$\int\int\left(\frac{\phi(x)-\phi(x)}{x-y}\right)^2 \mu(dx)\mu(dy) \geq CVar\mu(\phi) \tag{62}$$

For every $\phi \in C_0^1(\mathbf{R})$

As opposed to (61) which requires certain conditions on the measure μ_V, it turns out that (62) is always satisfied for any compactly supported measure μ with some constant. As was shown in [200] for the semicircular law, one can completely characterize the distribution in terms of the constant C.

After the use of convexity, inequality (61) may actually be interpreted as a spectral gap as follows. On $L^2\left(\frac{\mathrm{I}_{[-2,2]}(x)dx}{\sqrt{4-x^2}}\right)$ take the Jacobi operator

$$Lf = -\left(1-x^2\right)f''(x)+xf'(x)$$

and the counting number operator defined by

$$NT_n = nT_n$$

Where T_n are the Chebyshev polynomials of the first kind, which are orthogonal in $L^2\left(\frac{\mathrm{I}_{[-2,2]}(x)dx}{\sqrt{4-x^2}}\right)$. Then, (61) for $V(x)-px^2/2$ is equivalent to $L \geq N$.

Inequality (62) in the case of $V(x)-px^2/2$ can also be seen as the spectral gap for the counting number operator on $L^2\left(1_{[-2,2]}(x)\sqrt{4-x^2}dx\right)$ with respect to the basis given by the Chebyshev polynomials of second kind. A more general situation is discussed in this Section which includes both versions of the Poincaré inequalities.

Throughout this section we consider lower semicontinuous potentials $V:\mathbf{R}\to\mathbf{R}$ such that

$$\lim_{|x|\to\infty}\left(V(x)-2\log|x|\right)=\infty \tag{63}$$

For a given Borel set $\Gamma\subset\mathbf{R}$, denote by $\mathrm{P}(\Gamma)$ the set of probability measures supported on Γ.

The logarithmic energy with external potential V is defined by

$$\mathrm{E}_V(\mu):=\int V(x)\mu(dx)-\int\int\log|x-y|\mu(dx)\mu(dy)$$

whenever both integrals exist and have finite values. In particular for measures μ which have atoms, $\mathrm{E}_V(\mu)=+\infty$ because the second integral is +∞. It is known (see [52] or [196]) that under condition (63) there exists a unique minimizer of E_V in the set $\mathrm{P}(\mathbf{R})$ and the solution μ_V is compactly supported. The variational characterization of the minimizer μ_v ([52]) is that for a constant $C\in\mathbf{R}$,

$V(x)\geq 2\int\log|x-y|\,\mu_v(dy)+C$ for quasi-every $x\in\mathbf{R}$,

$$V(x)=2\int\log|x-y|\,\mu_V(dy)+C \text{ forquasi-every } x\in\operatorname{supp}(\mu_V) \tag{64}$$

where $\sup(\mu_v)$ stands for the support of μ. If μ is such that $\mathrm{E}_V(\mu)<\infty$, then Borel quasieverywhere sets have μ measure 0 and thus the properties above hold almost surely with respect to μ.

For simplicity of the notation, we will drop the subscript V from E_V unless the dependence of the potential has to be highlighted.

Now we summarize some known facts about the equilibrium measure and its support as one can easily deduce them from [52] and [196].

Theorem (6.3.1)[177]:

(i) Let V be a potential satisfying (63) and $\alpha \neq 0, \beta \in \mathbf{R}$. Set $V_{\alpha,\beta}(x) = V(\alpha x + \beta)$. Then, $\mu_{V_{\alpha,\beta}} = ((id - \beta)/\alpha)_{\#}\mu_V$ and

$$\mathrm{E}_V(\mu_V) = \mathrm{E}_{V_{\alpha,\beta}}\left(\mu_{V_{\alpha,\beta}}\right) - \log|x| \tag{65}$$

(ii) If V is convex satisfying (63), then the support of the equilibrium measure μ_v consists of one interval [a, b] where a and b solve the system

$$\begin{cases} \dfrac{1}{2\pi}\displaystyle\int_a^b V'(x)\sqrt{\dfrac{x-a}{b-x}}dx = 1 \\ \dfrac{1}{2\pi}\displaystyle\int_a^b V'(x)\sqrt{\dfrac{b-x}{x-a}}dx = -1 \end{cases} \tag{66}$$

(iii) Let V be either aC^2 satisfying (63) whose equilibrium measure has support [[a, b]. Then the equilibrium measure μ_V has density $g(x)$ given by

$$g(x) = 1_{[a,b]}\frac{\sqrt{(x-a)(b-x)}}{2\pi^2}\int_a^b \frac{V'(y) - V'(x)}{(y-x)\sqrt{(y-a)(b-y)}}dy, \tag{67}$$

(iv) If V is C^2, then

$$V'(x) = p.v\int \frac{2}{x-y}\mu_v(dx) \text{ for } \mu_V - a.s \text{ all } x \in \operatorname{supp}(\mu_V), \tag{68}$$

where p.v. stands for the principal value integral. Notice that the principal value makes sense as μ_V has a continuous density.

We mention as a basic example that if $V(x) = \rho x^2$ is quadratic, then μ_V is the semicircular law.

$$\mu_V(dx) = 1_{\left[-\sqrt{2/p}, \sqrt{2/p}\right]}(x)\sqrt{2\rho - \rho^2 x^2}\frac{dx}{\pi}$$

In this work, for $p \geq 1$, we use $\mathrm{W}_p(\mu, v)$ for the Wasserstein distance on the space of probability measures on $\mathbf{R}$ defined as

$$\mathrm{W}_p(\mu, \mathrm{V}) = \inf_{\pi \in \prod(\mu,v)}\left(\int\int |x-y|^p \pi(dx, dy)\right)^{\frac{1}{p}} \tag{69}$$

with $\Pi(\mu, \mathrm{V})$ denoting the set of probability measures on $\mathbf{R}^2$ with marginals μ and

$$\mathrm{W}_p^p(\mu, \mathrm{V}) = \int |\Theta(x) - x|^p \, \mathrm{V}(dx) \tag{70}$$

For a detailed discussion on this topic .

Our first result concerns the free version of the transportation cost inequality. As discussed in the introduction, the first assertion for strictly convex potentials was initially proved by large matrix approximation in [66]. The strategy of proof is inspired from [80, 32, 179] (see [102]).

Theorem (6.3.2) [177]: (Transportation inequality).

(i) If V is C^2 and $V(x) - \rho x^2$ is convex for some $\rho > 0$, then for any probability measure μ on $\mathbf{R}$,

$$\rho \mathrm{W}_2^2(\mu, \mu_V) \le \mathrm{E}(\mu) - \mathrm{E}(\mu_V) \tag{71}$$

If $V(x) - \rho x^2$, then the equality in (71) is attained for measures $\mu = \theta_{\#} \mu_{\mathrm{V}}$, with $\theta(x) = x + m$, therefore the constant ρ in front of $\mathrm{W}_2^2(\mu, \mu_V)$ is sharp.

(ii) Assume that V is C^2, convex and $V''(x) \ge p > 0$ for all $|x| \ge r$. Then, there is a constant $C = C(r, p, \mu_v, \mathrm{V}) > 0$, such that

$$C\mathrm{W}_2^2(\mu, \mu_V) \le \mathrm{E}(\mu) - \mathrm{E}(\mu_V) \tag{72}$$

(iii) In the case V is C^2 and $V(x) - \rho|x|^p$ is convex for some real number $\rho > 0$, then, for any probability measure μ on $\mathbf{R}$,

$$c_p \rho \mathrm{W}_p^p(\mu, \mu_V) \le \mathrm{E}(\mu) - \mathrm{E}(\mu_V) \tag{73}$$

Where $C_p = \inf_{x \in \mathbf{R}} \left(|1 + x|^p - |x|^p - p\,sign(x)|x|^{p-1} \right) > 0$.

Proof: (i) Since there is nothing to prove in the case $\mathrm{E}(\mu) = \infty$, we assume that $\mathrm{E}(\mu) < \infty$. In this case we also have that the measure μ and μ_V both have second finite moments.

Now we take the non-decreasing transportation map θ such that $\theta_{\#} \mu_V = \mu$ which exists due to the lack of atoms of μ_V. Using the transport map θ, we first write

$$\mathrm{E}(\mu)-\mathrm{E}(\mu_V)=\int\Big(V(\theta(x))-V(x)-V'(x)(\theta(x)-x)\Big)\mu_V(dx)$$
$$+\int\left(\frac{\theta(x)-\theta(y)}{x-y}-1-\log\frac{\theta(x)-\theta(y)}{x-y}\right)\mu_V(dx)\mu_V(dy) \qquad (74)$$

where in between we used the variational equation (68) to justify that

$$\int V'(x)(\theta(x))\mu_V(dx)=2\int\int\frac{\theta(x)-x}{x-y}\mu_V(dy)\mu_V(dx)$$
$$=\int\frac{(\theta(x)-x)-(\theta(y)-y)}{x-y}\mu_V(dy)\mu_V(dx)$$

Since $V(x)-\rho x^2$ is convex, for any x, y the following holds

$$V(y)-V(x)-V'(x)(y-x)\geq\rho\left(y^2-x^2-2x(y-x)\right)=\rho(y-x)^2$$

On the other hand since $a-1\geq\log(a)$ for any $a\geq 0$, Eqs. (74) and (70) yield (71).

In the case $V(x)-\rho x^2$ it is easy to see that for $\theta(x)=x+m$, all inequalities involved become equalities, thus we attain equality in (71) for translations of μ_V.

(ii) We start the proof with (74), whereas this time we need to exploit the logarithmic term to get our inequality. The idea is to use the strong convexity where $\psi(x):=\theta(x)-x$ takes large values and for small values of $\psi(x)$ we try to compensate this with the second integral of (74).

Notice in the first place that by Taylor's theorem we have that

$$V(y)-V(x)-V'(x)(y-x)=(y-x)^2\int_0^1 V''\big((1-\tau)x+\tau y\big)(1-\tau)d\tau \qquad (75)$$

Now, let us assume that the support of the equilibrium measure μ_v is $[\mathrm{a},\mathrm{b}]$. Next, $V''(x)\geq 0$ and $V''(x)\geq p$ for $|x|\geq r$, implies that for $|y|\geq 2r+2\max\{|a|,|b|\}$, we obtain that

$$V(y)-V(x)-V'(x)(y-x)\geq(y-x)^2\int_{1/2}^1 V''\big((1-\tau)x+\tau y\big)(1-\tau)d\tau$$
$$\geq p(y-x)^{2/8}\ \textit{for any}\ x\in[a,b]$$

Now write $\theta(x)=x+\psi(x)$. Thus using (74), and denoting

$\mathbf{R}=2r+2\max\{|a|,|b|\}$ we continue with

$$\int\left(V\left(\theta(x)\right)-V(x)-V'(x)\left(\theta(x)-x\right)\right)\mu_V(dx)$$
$$\geq\frac{1}{2}\int\psi^2(x)\int_0^1 V''\left(x+\tau\psi(x)\right)(1-\tau)d\tau\mu_V(dx)\geq\frac{p}{16}\int_{|\psi|\geq\mathbf{R}}\psi^2(x)\mu_V(dx) \tag{76}$$

This inequality provides a lower bound of the first term in (74). Further, it is not hard to check that

$$\begin{aligned}&\int_{|\psi|\geq\mathbf{R}}\psi^2(x)\mu_V(dx)\\&=\frac{1}{2}\int 1_{|\psi|\geq\mathbf{R}}(x)\psi^2(x)\mu_V(dx)+\frac{1}{2}\int 1_{|\psi|\geq\mathbf{R}}(y)\psi^2(y)\mu_V(dy)\\&\geq\frac{1}{8}\int\int 1_{|\psi(x)-\psi(y)|\geq 2\mathbf{R}}(x,y)\left|\psi(x)-\psi(y)\right|^2\mu_V(dx)\mu_V(dy)\end{aligned} \tag{77}$$

Now we treat the second integral on the left-hand side of (74). Use that $t-\log(1+t)\geq|t|-\log(1+|t|)$ for any $t>-1$ together with the fact that $t-\log(1+t)$ is an increasing function for $t\geq 0$ to argue that

$$\begin{aligned}&\int\int\left(\frac{\psi(x)-\psi(y)}{x-y}-\log\left(1+\frac{\psi(x)-\psi(y)}{x-y}\right)\right)\mu_V(dx)\mu_V(dy)\\&\geq\int\int\left(\frac{|\psi(x)-\psi(y)|}{b-a}-\log\left(1+\frac{|\psi(x)-\psi(y)|}{b-a}\right)\right)\mu_V(dx)\mu_V(dy)\end{aligned} \tag{78}$$

Further, for $s\geq 0$ and $u,v>0$ we have

$$us^2+s-\log(1+s)\geq\begin{cases}\dfrac{v-\log(1+v)}{v^2}s^2, & 0\leq s\leq v\\ us^2 & ,v\leq s\end{cases}\geq\min\left\{u,\frac{v-\log(1+v)}{v^2}\right\}s^2$$

This inequality used for $u=\dfrac{\rho(b-a)^2}{128}$ and $v=\dfrac{2R}{b-a}$ in combination with (77) and (78) yields for the choice of $C=\min\left\{u,\left(v-\log(1+v)\right)/v^2\right\}$ that

$$\begin{aligned}&\frac{\rho}{16}\int_{|\psi|\geq\mathbf{R}}\psi^2(x)\mu_V(dx)+\int\int\left(\frac{\psi(x)-\psi(y)}{x-y}-\log\left(1+\frac{\psi(x)-\psi(y)}{x-y}\right)\right)\mu_V(dx)\mu_V(dy)\\&\geq c\int\int\left(\psi(x)-\psi(y)\right)^2\mu_V(dx)\mu_V(dy)=c\left[\int\psi^2(x)\mu_V(dx)-\left(\int\psi(x)\mu_V(dx)\right)^2\right]\end{aligned} \tag{79}$$

This shows that $\mathrm{E}(\mu)-\mathrm{E}(\mu_V)$ is bounded below by a constant times the variance of ψ. Notice that $\mathrm{W}_2^2(\mu,\mu_V)=\int\psi^2(x)\mu_V(dx)$ and in order to complete the proof we have to replace the variance of ψ by the integral of ψ^2 with respect to μ_V. This boils down to estimating the μ_V integral of ψ in terms of the integral of ψ^2.

To this end, use Cauchy's inequality:

$$\left(\int\psi(x)\mu_V(dx)\right)^2\leq\int\psi^2(x)\left(1+\frac{1}{2c}\int_0^1 V''\left(x+\tau\psi(x)\right)(1-\tau)\,d\tau\right)\mu_V(dx)$$
$$\times\int\frac{1}{1+\frac{1}{2c}\int_0^1 V''\left(x+\tau\psi(x)\right)(1-\tau)\,d\tau}\mu_V(dx)$$

This inequality combined with Eqs. (74), (76) and (79), results with

$$\mathrm{E}(\mu)-\mathrm{E}(\mu_V)\geq\int\psi^2(x)\left(C+\frac{1}{2}\int_0^1 V''\left(x+\tau\psi(x)\right)(1-\tau)\,d\tau\right)\mu_V(dx)$$
$$\times\int\frac{\int_0^1 V''\left(x+\tau\psi(x)\right)(1-\tau)\,d\tau}{2c+\int_0^1 V''\left(x+\tau\psi(x)\right)(1-\tau)\,d\tau}\mu_V(dx)$$
$$\geq\int\frac{\int_0^1 V''\left(x+\tau\psi(x)\right)(1-\tau)\,d\tau}{2c+\int_0^1 V''\left(x+\tau\psi(x)\right)(1-\tau)\,d\tau}\mu_V(dx)\mathrm{W}_2^2(\mu,\mu_V)$$

where here we used the convexity encoded into $V''\geq 0$ and the fact that

$\mathrm{W}_2^2(\mu,\mu_V)=\int\psi^2(x)\mu_V(dx)$ to get the lower bound of the first integral.

From the previous inequality, it becomes clear that we are done as soon as we prove that the quantity in front of $\mathrm{W}_2^2(\mu,\mu_V)$ is bounded from below by a positive constant uniformly in ψ. To carry this out, notice that V'' can not be identically zero on [a, b]. Indeed, if V'' were identically zero on [a, b], then we would have that $V'(x)=K$ for all $x\in[a,b]$, and this plugged into Eq. (66), yields that $K(a,b)=2$ and $K(a,b)=-2$, a system without a solution. Therefore V'' is not identically 0 on $[a,b]$. If $\left|\psi(x)\right|>\mathbf{R}$, then $V''\left(x+\tau\psi(x)\right)\geq\rho$ for

$\frac{1}{2} \le \tau \le 1$, which implies $\int_0^1 V''(x+\tau\psi(x)) \times (1-\tau) \ge \rho/8$. On the other hand, if $|\psi(x)| \le \mathbf{R}$, then

$$\int_0^1 V''(x+\tau\psi(x))(1-\tau)\,d\tau \ge \int_0^\delta V''(x+\tau\psi(x))(1-\tau)\,d\tau \ge \frac{\delta}{2} \inf_{|y-x|\le\delta R} V''(y)$$

for all $0 \le \delta \le 1$. Define

$$w(x) = \sup_{\delta\in[0,1]} \min\left\{\frac{\rho}{8}, \frac{\delta}{2} \inf_{|y-x|\le\delta R} V''(y)\right\}$$

Since V'' is not identically 0 on $[a, b]$, it follows that w is not identically zero on $[a, b]$. With this we obtain that

$$\int_0^1 V''(x+\tau\psi(x))(1-\tau)\,d\tau \ge w(x) \ge 0$$

and then that

$$c\int \frac{\int_0^1 V''(x+\tau\psi(x))(1-\tau)\,d\tau}{2c+\int_0^1 V''(x+\tau\psi(x))(1-\tau)\,d\tau}\mu_V(dx) \ge C = \int \frac{cw(x)}{2c+w(x)}\mu_V(dx) > 0$$

which finishes the proof of (72) with this choice of C.

(iii) For the inequality (73), we follow the same route as in the proof of (71), the only change this time being that $V(x)-\rho|x|^p$ is convex, and thus we obtain

$$V(y)-V(x)-V'(x)(y-x) \ge \rho\left(|y|^p - |x|^p - p\,sign(x)|x|^{p-1}(y-x)\right) \tag{80}$$

Writing $\theta(x) = x+\psi(x)$, and using (74) together with $a-1 \ge \log(a)$ for $a \ge 0$, one arrives at

$$\mathrm{E}(\mu) - \mathrm{E}(\mu_V) \ge \rho\int\left(|x+\psi(x)|^p - |x|^p - p\,sign(x)|x|^{p-1}\psi(x)\right)\mu_V(dx)$$

Now we use the fact that for all $a, b \in \mathbf{R}$,

$$|a+b|^p - |b|^p - p\,sign(b)|b|^{p-1}a \ge Cp|a|^p \tag{81}$$

which applied to the above inequality in conjunction to (70), yields inequality (73).

Remark (6.3.3) [177]:(i) The C^2 regularity of V for (71) can be dropped (see [102]) but to simplify the presentation here we decided to consider only this case.

(ii) If $V(x)-\rho|x|^p$ is convex, then using inequalities (73), (72) and Young's inequality we obtain that for any $2 \le K \le \rho$, there exists a constant $c = c(k, p, \rho, \mu_v, V)$ such that

$$cW_k^k(\mu, \mu_V) \le E(\mu) - E(\mu_V)$$

(iii) We want to point out that the inequalities (73) and (72) are somehow complementary to each other. For example, if we take $V(x)-\rho|x|^p$ with $p>1$ and the measure $\mu = \theta_{\#}\mu_v$ for $\theta(x) = x+m$, then Eq. (73) takes the form

$$c_p m^p \le \int \left(|x+m|^p - |x|^p\right) \mu_V(dx) \tag{82}$$

while Eq. (72) becomes

$$Cm^2 \le \int \left(|x+m|^p - |x|^p\right) \mu_V(dx)$$

which, because it is easy to check that μ_v is symmetric, is the same as

$$Cm^2 \le \int \left(|x+m|^p - |x|^p - p\,sign(x)|x|^{p-1} m\right) \mu_V(dx) \tag{83}$$

Notice here that (82) is in the right scale for large m as (83) is in the right scale for m close to 0, because in this case the integrand is of the size m^2. It seems that Talagrand's transportation inequality in this context has two a spects, one is the large $W_p(\mu, \mu_v)$ which is dictated by the potential V for large values and results with Eq. (73) and the small $W_2(\mu, \mu_V)$ regime which is dictated by the repulsion effect of the logarithm and results with Eq. (72).

(iv) It is not clear whether inequality (72) still holds for the case of a potential V which is not convex. Of interest would be the particular case $V(x) = ax^4 + bx^2$ for some $a > 0$ and $b < 0$. This example actually raises the question of the stability of transportation inequality under bounded perturbations.

(v) Very likely the constant C_p in (73) is not sharp.

In this section, we investigate some potential independent transportation inequalities. A transportation inequality in the form of (72) can not possibly hold without a quadratic growth at infinity. Also, the proof of (72) might lead to the conclusion that the logarithmic term plays a more important role. Therefore the natural question one may ask is whether there is a manifestation of this fact in some sort of transportation type inequality which is independent of the potential involved. The main question reduces to hint some appropriate distance one needs to use to replace the Wasserstein distance in Theorem (6.3.3).We investigate in this section several possibilities, starting with the free version of the classical Pinsker's inequality.

The Pinsker's inequality classically states that ([96] and [126])

$2\|\mu - \mathrm{V}\|_V^2 \leq H(\mu|\mathrm{V})$ for any μ, V probability measures on $\mathbf{R}$,

where $\|\mu - \mathrm{V}\|_{\mathrm{V}}$ is the total variation distance between μ and V and $H(\mu|\mathrm{V})$ is the relative entropy between μ and V. This in particular shows that if μ_n convergence to μ in entropy, then μ_n converges to μ is a very strong sense.

The same natural question can be posed in the logarithmic entropy context. For a given potential V , is there an inequality of the form

$$C\|\mu - \mu_V\|_V^2 \leq \mathrm{E}(\mu) - (\mu_V)$$

for a given constant $C > 0$ and any probability distribution μ on $\mathbf{R}$

It turns out that these inequalities do not hold for the logarithmic energy. In fact, we will show that even a weaker inequality of the form

$$C\left\|F_\mu - F_{\mu_V}\right\|_u^2 \leq \mathrm{E}(\mu) - \mathrm{E}(\mu_V) \tag{84}$$

does not hold, where F_μ denotes the cumulative function of a probability measure μ on the line. Even though the uniform distance does not have the same widespread use in probability it appears for example in the Berry–Esseen type estimates for the convergence in the central limit theorem. This is the reason why we consider this distance as the first next best candidate wherever the total variation fails. Clearly this metric gives a stronger topology as the topology of weak convergence.

We will construct a counterexample to (84) in the case of $V(x) = 2x^2$, for which the equilibrium measure is

$$\mu_V(dx) = 1_{[-1,1]}(x)\frac{2\sqrt{1-x^2}}{\pi}dx$$

the semicircular law on $[-1,\ 1]$. Consider now the sequence

$$\mu_V(dx) = 1_{[-1,1]}(x)\frac{2\sqrt{1-x^2}}{\pi}dx + \frac{\sum_{k=2}^{2n-1}(-1)^k T_{2k+1}(x)}{4(n^2-1)\pi\sqrt{1-x^2}}dx$$

where T_k is the kth Chebyshev polynomial of the first kind. With these choices we have that

$$\mathrm{E}(\mu_n) - \mathrm{E}(\mu_V) \le \frac{\pi^2}{\log(n/3)}\left|F_{\mu n} - F_{\mu V}\right|_u^2 \text{ for all } n \ge 4. \tag{85}$$

Let us point out that μ_n is indeed a probability measure. This requires a little proof but it is entirely elementary.

To prove (84), notice that since the support of μ_n is the same as the support of μ_v, we have from (64) that

$$\mathrm{E}(\mu_n) - \mathrm{E}(\mu_V) = -\int\int \log|x-y|(\mu_n - \mu_V)(dx)(\mu_n - \mu_V)(dy) \tag{86}$$

Next remark that $\mu_n = \cos_{\#}(f_n\lambda)$ and $\mu_v = \cos_{\#}(g\lambda)$, where λ is the Lebesgue measure on $[0,\pi]$ and

$$f_n(t) = \frac{1-\cos(2t)}{\pi} + \frac{1}{4\pi(n^2-1)}\sum_{k=2}^{2n-1}(-1)^k\cos((2k+1)t), \quad g(t) = \frac{1-\cos(2t)}{\pi}$$

and further

$$-\int\int \log|x-y|(\mu_n - \mu_V)(dx)(\mu_n - \mu_V)(dy) = -\int_0^\pi\int_0^\pi \log|\cos t - \cos s|\, h_n(t)h_n(s)\,dtds$$

where $h_n = f_n - g$.

Now we provide a formula for the logarithmic energy we learnt from [260] and have not seen it elsewhere. Here is a quick description. Write first

$\cos t = \left(e^{it} + e^{-it}\right)/2$ and $\cos s = \left(e^{it} + e^{-it}\right)/2$ so $|\cos t - \cos s| = \left|\left(e^{it} + e^{-it}\right)/2 - \left(e^{is} + e^{-is}\right)\right|/2 = \left|1 - e^{i(t+s)}\right|\left|1 - e^{i(t-s)}\right|/2$ and so, for $t \ne s$, and t or s not equal to π,

$$\log\left|\cos t-\cos s\right| = -\log 2+\operatorname{Re}\left(\log\left(1-e^{i(t+s)}\right)+\log\right)\left(1-e^{i(t-s)}\right)$$
$$= \log 2-\sum_{t=1}^{\infty}\operatorname{Re}\left(e^{i\ell(t+s)}/\ell+e^{i\ell(t-s)}\ell\right) = \log 2-\sum_{\ell=1}^{\infty}\frac{2}{\ell}\cos(\ell t)\cos(\ell s)$$

From this, one gets to

$$-\int_0^{\pi}\int_0^{\pi}\log\left|\cos t-\cos s\right|h_n(t)h_n(s)\,dtds = \frac{2}{\ell}\left(\int_0^{\pi}\cos(\ell t)h_n(t)\,dt\right)^2 \tag{87}$$

But now,

$$\int_0^{\pi}\cos(\ell t)h_n(t)\,dt = \frac{1}{4\pi\left(n^2-1\right)}\sum_{k=2}^{2n-1}(-1)^k\int_0^{\pi}\cos(\ell t)\cos((2k+1)t)\,dt$$
$$= \begin{cases}\dfrac{(-1)^{(t-1)/2}}{8\left(n^2-1\right)}, & 4\le\ell\le 4n \textit{ and odd}\\ 0 & \textit{otherwise}\end{cases}$$

and thus

$$\begin{aligned}-\int_0^{\pi}\int_0^{\pi}\log\left|\cos t-\cos s\right|h_n(t)h_n(s)\,dtds &= \sum_{t=1}^{\infty}\frac{2}{\ell}\left(\int_0^{\pi}\cos(\ell t)h_n(t)\,dt\right)^2\\ &= \frac{1}{32\left(n^2-1\right)^2}\sum_{t=2}^{2n-1}\frac{1}{2\ell+1}\end{aligned} \tag{88}$$

On the other hand $\left|F_{\mu n}-F_{\mu v}\right|_u = \left|F_{fn\lambda}-F_{g\lambda}\right|_u = \sup_{x\in[0,\pi]}\left|\int_0^x h_n(t)dt\right|$ and

$$\int_0^x h_n(t)dt = \frac{1}{4\pi\left(n^2-1\right)}\sum_{\ell=2}^{2n-1}\frac{(-1)^{\ell}\sin((2\ell+1)x)}{2\ell+1}$$

from which for $x=\pi/4$, we obtain

$$\left|F_{\mu n}-F_{\mu v}\right|_u = \sup_{x\in[0,\pi]}\left|\int_0^x h_n(t)dt\right| \ge \frac{1}{4\pi\left(n^2-1\right)}\sum_{\ell=2}^{2n-1}\frac{1}{2\ell+1} \tag{89}$$

Combining (88) and (89) we get

$$\frac{\pi^2}{2\sum_{\ell=2}^{2n-1}\frac{1}{2\ell+1}}\left|F_{\mu n}-F_{\mu v}\right|_u^2 \ge -\int\int\log\left|x-y\right|(\mu_n-\mu_v)(dx)(\mu_n-\mu_v)(dy) \tag{90}$$

which together with the fact that $\sum_{\ell=2}^{2n-1}\frac{1}{2\ell+1}\geq\frac{1}{2}\log(n/3)$ for $n\geq 4$ and (86), we finally arrive at (85).

The example shown above has the property that $\mathrm{E}(\mu_n)-\mathrm{E}(\mu_v)$ converges to 0 when n goes to infinity, and also that $\left|F_{\mu n}-F_{\mu v}\right|_u$ converges to zero. Despite the fact that (84) does not hold, we will see below in Corollary (6.3.11) that if $\mathrm{E}(\mu_n)-\mathrm{E}(\mu_v)$

converges to 0, then $\left|F_{\mu n}-F_{\mu v}\right|_u$ always converges to 0.

We consider now a weak form of (84). To do this we define the distance

$$d(\mu,\mathrm{V})=\sup_{a,b\in\mathbf{R}}\left|\int e^{-|ax+b|}\mu(dx)-\int e^{-|ax+b|}\mathrm{V}(dx)\right| \tag{91}$$

With this definition we have the following result.

Theorem (6.3.4) [177]: For any potential V satisfying (63), we have that for any compactly supported measure μ,

$$4\pi^3 d^2(\mu,\mu_V)\leq\mathrm{E}(\mu)-\mathrm{E}(\mu_V) \tag{92}$$

Proof: Using Eqs. (63) and (64), we get for any compactly supported measure μ with $\mathrm{E}(\mu)$ finite,

$$\mathrm{E}(\mu)-\mathrm{E}(\mu_V)\geq-\int\int\log|x-y|(\mu-\mu_V)(dx)(\mu-\mu_V)(dy)$$

We will prove that for any measures μ and V with compact support such that

$-\int\int\log|x-y|\mu(dx)\mu(dy)<\infty$ and $-\int\int\log|x-y|\mathrm{V}(dx)\mathrm{V}(dy)<\infty$, we have that

$$4\pi^3 d^2(\mu,\mu_V)\leq\int\int\log|x-y|(\mu,\mathrm{V})(dx)(\mu,\mathrm{V})(dy) \tag{93}$$

which shows that (93) implies (92).

Now we use [196] to write

$$-\int\int\log|x-y|(\mu-\mu_V)(dx)(\mu-\mu_v)(dy)=\int_0^\infty\frac{|\hat{\mu}(t)-|\hat{\mu}_V(t)|}{t}dt$$

where the hat stands for the Fourier transform, and continue with

$$\int_0^{\infty}\frac{|\hat{\mu}(t)-\hat{v}(t)|^2}{t}dt=\frac{1}{2}\int_0^{\infty}\frac{|\hat{\mu}(t)-\hat{v}(t)|^2}{|t|}dt\geq|a|\int_{-\infty}^{\infty}\frac{|\hat{\mu}(t)-\hat{v}(t)|^2}{a^2+t^2}dt$$

$$\geq\frac{a^2}{\pi}\left|\int_{-\infty}^{\infty}\frac{\hat{\mu}(t)-\hat{v}(t)}{a^2+t^2}e^{-ict}dt\right|^2$$

for any $ac\in\mathbf{R}$ with $a\neq 0$. Further, using the inversion formula for the Fourier transform, one has

$$\int_{-\infty}^{\infty}\frac{\hat{\mu}(t)-\hat{\mathrm{V}}(t)}{a^2+t^2}e^{-ict}dt=2\pi\int\hat{\phi}(x)(\mu-\mathrm{V})(dx)=\frac{2\pi^2}{|a|}\int e^{-|ax+b|}(\mu-\mathrm{V})(dx) \qquad (94)$$

Because for

$$\phi(t)=\frac{e^{ict}}{a^2+t^2},$$

$$\hat{\phi}(x)=\int\frac{e^{i(x+c)t}}{a^2+t^2}dt=\frac{\pi e^{-|a(x+t)|}}{|a|} \qquad (95)$$

The next result is collecting facts about how strong the topology induced by d is.

Proposition (6.3.5) [177]:

(i) d is a distance on $p(\mathbf{R})$ and if $d(\mu_n,\mu)\to 0$, then $\mu_n\to_{n\to\infty}\mu$ in the weak topology. In addition $d(\delta_a,\delta_b)=1$ for $a\neq b$, thus the topology induced by d is strictly stronger than the weak convergence topology.

(ii) For any two probability measures μ and V,

$$d(\mu,\mathrm{V})\leq 2|F_{\mu}-F_{\mathrm{V}}|_u \qquad (96)$$

(iii) If V satisfies condition (63), then $\mathrm{E}_{\mathrm{V}}(\mu_n)\xrightarrow[n\to\infty]{}\mathrm{E}_{\mathrm{V}}(\mu_{\mathrm{V}})$

implies $|F_{\mu n}-F_{\mu\mathrm{V}}|_u\xrightarrow[n\to\infty]{}0$

Proof: (i) To prove that d is a distance the only non trivial fact is that for two probability measures μ and $\mathrm{V}, d(\mu,\mathrm{V})=0$ implies $\mu=\mathrm{V}$. Thus from Eq. (95), we obtain for $a=1$ that for all $c\in\mathbf{R}$,

$$\int_{-\infty}^{\infty}\frac{(\hat{\mu}(t)-\hat{\mathrm{V}}(t))e^{-ict}}{1+t^2}dt=0$$

Since this holds true for any $c \in \mathbf{R}$, it implies that the Fourier transform of the function $t \to \frac{\hat{\mu}(t) - \hat{\mathrm{V}}(t)}{1+t^2}$ is 0, which means that the function in discussion must be 0. This means that $\hat{\mu} = \hat{v}$, or equivalently that $\mu = \mathrm{V}$.

Let $\mathrm{L}(\mu, \mathrm{V})$ stand for the Levy distance which induces the weak topology on $\mathrm{P}(\mathbf{R})$. Let $d(\mu_n, \mu) \xrightarrow[n\to\infty]{} 0$. Assume now that there exists $\varepsilon > 0$ and a subsequence such that $\mathrm{L}\left(\mu_{n_k}, \mu\right) \geq \varepsilon$. Otherwise said, the sequence μ_n has a subsequence which is not convergent to μ. Since, we are dealing with probability measures, there is a subsequence μ_{n_k} which is vaguely convergent to a measure V with total mass less than 1. This means that for any continuous function ϕ which is vanishing at infinity, we have that

$$\int \phi d\mu_{n_{k_l}} \to \int \phi d\mathrm{V}$$

We can apply this for functions $\phi(x) = e^{-|ax+b|}$ where $a \neq 0$ and infer that

$$\int e^{-|ax+b|} \mu_{n_{k_l}}(dx) \xrightarrow[l\to\infty]{} \int e^{-|ax+b|} \mathrm{V}(dx) \text{ for all } a \neq 0,\ b \in \mathbf{R}.$$

On the other hand, because $d\left(\mu_{n_{k_l}}, \mu\right) \xrightarrow[l\to\infty]{} 0$, these considerations result with

$$\int e^{-|ax+b|} \mu(dx) = \int e^{-|ax+b|} \mathrm{V}(dx) \text{ for all } a \neq 0,\ b \in \mathbf{R}.$$

Further, using the dominated convergence for $\mathrm{b} = 0$ and $\mathrm{a} \to 0$, we obtain that V is a probability measure. From the discussion at the beginning of this proof, it also follows that $\mathrm{V} = \mu$ and this in turn results with $\mu_{n_{k_l}}$ being weakly convergent to μ, a contradiction. This proves that the convergence in the metric d implies weak convergence.

It is obvious that $d(\mu, \mathrm{V}) \leq 1$ for any measures μ and V. For the case of discrete measures, we also have that

$1 \geq d(\delta_a, \delta_b) \geq \int e^{-\alpha|x-a|} \delta_a(dx) - \int e^{-\alpha|x-a|} \delta_b(dx)$ for any $\alpha > 0$, which yields that $1 \geq d(\delta_a, \delta_b) \geq 1 - e^{-\alpha|b-a|}$ for all $\alpha > 0$. Letting $\alpha \to \infty$, we get that

$d(\delta_a, \delta_b) = 1$ for $a \neq b$ which shows that convergence in d is strictly stronger than convergence in the weak topology.

(ii) From the fact that for any finite positive measure μ,

$$\int_{(0,\infty)} \left(1-e^{-\alpha y}\right)\mu(dx) = \int_{(0,\infty)} \alpha e^{-\alpha y}\mu\left((y,\infty)\right)dy$$

we deduce that

$$\int e^{-\alpha|x-a|}(\mu-\mathrm{V})(dx) = \int_{(0,\infty)} \alpha e^{-\alpha y}\left[F_\mu(a-y)-F_\mu(a+y)-F_{\mathrm{V}}(a-y)+F_{\mathrm{V}}(a+y)\right]dx$$

which easily yields (96).

(iii) We actually show that if μ_n and μ are compactly supported probability measures such that

$$\iint \log|x-y|\,\mu(dx)\mu(dy) < \infty, \qquad -\iint \log|x-y|\,\mu_n(dx)\mu_n(dy) < \infty$$

and

$$\lim_{n\to\infty}\iint \log|x-y|\left(\mu_n-\mu\right)(dx)\left(\mu_n-\mu\right)(dy) = 0$$

then $\left|F_{\mu^n}-F_\mu\right|_u \xrightarrow[n\to\infty]{} 0$. From (72) and the first part, we obtain that μ_n converges weakly to μ. In addition, none of the measures μ_n or μ have atoms. Thus $F_{\mu n}$ and F_μ are continuous functions which combined with the weak convergence implies that $F_{\mu n}$ converges pointwise to F_μ. Since the functions $F_{\mu n}$ and F_μ are distributions of probability measures, it is an easy matter to check that the convergence is actually uniform.

Remark (6.3.6) [177]: We do not know if the topology of convergence in d is the same as the one defined by the metric $\left|F_\mu - F_{\mathrm{V}}\right|_u$.

This result might leave one wondering if a stronger convergence takes place. In other words, is it true that $\mathrm{E}_v(\mu_n) \xrightarrow[n\to\infty]{} \mathrm{E}_v(\mu_v)$, implies

$\left\|\mu_n-\mu\right\|_v \xrightarrow[n\to\infty]{} 0$ To this end, we can consider $V(x)=\log\left|\frac{|x|+\sqrt{x^2-1}}{2}\right|$

and notice (see [52]) that μ_v is the arcsine law of $[-1,\ 1]$. Thus if we consider

$$\mu_V(dx) = 1_{[-1,1]}(x)\frac{dx}{\pi\sqrt{1-x^2}}, \quad \mu_n(dx) = 1_{[-1,1]}(x)\frac{\left(1-T_n(x)\right)dx}{\pi\sqrt{1-x^2}},$$

then, using the same argument which led us to (87), with h_n there replaced by $h_n(x`)=\cos(nx)$ here, one arrives at $\mathrm{E}(\mu_n)-\mathrm{E}(\mu_v)=\frac{1}{n}$ while the total variation distance is $\|\mu_n-\mu\|_v \geq 1/4$.

We develop similarly the mass transportation method to prove the Log-Sobolev inequality in the free context. We define inspired by Voiculescu [45], the relative free Fisher information as

$$I(\mu)=\int(H\mu(x)-V'(x))^2\,\mu(dx) \text{ with } H\mu(x)=p.v\int\frac{2}{x-y}\mu(dy) \qquad (97)$$

for measures μ on which have density $\rho=d\mu/dx$ in $L^3(\mathbf{R})$. In this case the principal value integral is a function in L^3. Otherwise we let $\mathrm{I}(\mu)$ be equal to + ∞.

Theorem (6.3.7) [177]: (Log-Sobolev).

(i) If V is C^2 and $V(x)-\rho x^2$ is convex for some $\rho>0$, then for any probability measure μ on $\mathbf{R}$,

$$\mathrm{E}(\mu)-\mathrm{E}(\mu_V)\leq\frac{1}{4\rho}\mathrm{1}(\mu) \qquad (98)$$

Equality is attained for the case $V(x)-\rho x^2$ and $\mu=\theta_{\#}\,\mu_V$, where $\theta(x)=x+m$. Thus the inequality (98) is sharp for translations of μ_V.

(ii) If V is C^2 and $V(x)-\rho|x|^p$ is convex for some $\rho>0$ and $\rho>1$, then for any probability measure μ on $\mathbf{R}$,

$$\mathrm{E}(\mu)-\mathrm{E}(\mu_v)\leq\frac{K_p}{\rho^{q/p}}\mathrm{1}_q(\mu) \text{ where } I_q(\mu)=\int|H_\mu(x)-V'(x)|^q\times\mu(dx) \qquad (99)$$

where here q is the conjugate of p i.e. $1/q+1/p=1$ and the constant $k_p=(pc_p)^{q/p}$, with c_p from (73).

Proof:(i) We will assume that the measure μ has a smooth compactly supported density as the general case follows via approximation arguments discussed in details in [66]. Take the (increasing) transport map θ from μ_V into μ. We write the inequality (98) in the following equivalent way

$$
\begin{aligned}
&\frac{1}{4\rho}\int\left(H\mu\left(\theta(x)\right)-V'\left(\theta(x)\right)\right)^2\mu_V(dx)\\
&\quad+\int\left(V(x)-V\left(\theta(x)\right)-V'\left(\theta(x)\right)\left(x-\theta(x)\right)\right)\mu_V(dx)\\
&\quad-\int\left(H\mu\left(\theta(x)\right)-V'\left(\theta(x)\right)\right)\left(x-\theta(x)\right)\mu_V(dx)\\
&\quad+\int H\mu\left(\theta(x)\right)\left(x-\theta(x)\right)\mu_V(dx)-\int\int\log\frac{x-y}{\theta(x)-\theta(y)}\mu_V(dx)\mu_V(dy)\ge 0
\end{aligned}
\tag{100}
$$

Notice now that from the convexity of $V(x)-\rho x^2$, one obtains that

$$
\begin{aligned}
&V(x)-V\left(\theta(x)\right)-V'\left(\theta(x)\right)\left(x-\theta(x)\right)\\
&\ge\rho\left(x^2-\theta(x)^2-2\theta(x)\left(x-\theta(x)\right)\right)=\rho\left(x^2-\theta(x)\right)^2
\end{aligned}
\tag{101}
$$

Now,

$$
\begin{aligned}
&H\mu\left(\theta(x)\right)\left(x-\theta(x)\right)\mu_V(dx)\\
&=\int\left(x-\theta(x)\right)\int\frac{2}{\theta(x)-\theta(y)}\mu_V(dy)\mu_V(dx)=\int\int\left(\frac{x-y}{\theta(x)-\theta(y)}-1\right)\mu_V(dx)\mu_V(dy)
\end{aligned}
\tag{102}
$$

where one has to interpret the second integral here in the principal value sense, however since θ is increasing, the last integral is actually taken in the Lebesgue sense.

Using these, Eq. (100) may be rewritten as

$$
\begin{aligned}
&\frac{1}{4\rho}\int\left[H\mu\left(\theta(x)\right)-V'\left(\theta(x)\right)-2\rho\left(x-\theta(x)\right)\right]^2\mu_V(dx)\\
&\quad+\int\int\frac{x-y}{\theta(x)-\theta(y)}-1-\log\frac{x-y}{\theta(x)-\theta(y)}\mu_v(dx)\mu_V(dy)\ge 0
\end{aligned}
$$

which is seen to hold since $u-1-\log(u)\ge 0$ for $u\ge 0$.

Equality is attained for the case $V(x)-\rho x^2$ and $\theta(x)=x+c$, which corresponds to the translations of the measure μ_V.

(ii) With the same arguments used in the above proof and the proof of Theorem (6.3.2), we use Eqs. (80) and (81) to argue that

$$\frac{K_p}{\rho^{q/p}}\left|\int H\mu(x)-V'(x)\right|^q \mu(dx)-\mathrm{E}(\mu)-\mathrm{E}(\mu_V)$$

$$\geq \int\left[\frac{K_p}{\rho^{q/p}}\left|H\mu(\theta(x))-V'(\theta(x))\right|^q\right.$$

$$\left.+\left(V'(\theta(x))-H\mu(x)\right)\left(x-\theta(x)\right)+c_p\rho\left|x-\theta(x)\right|^p\right]\mu_V(dx)$$

$$+\int\int\frac{x-y}{\theta(x)-\theta(y)}-1\log\frac{x-y}{\theta(x)-\theta(y)}\mu_V(dx)\mu_V(dy)\geq 0$$

where we used Young's inequality $a^q/q+b^p/p\geq ab$ for $a,b\geq 0$ and the constant $K_p=\left(pc_p\right)^{q/p}q$

We devoted to the free analogue of the HWI inequality of Otto and Villani [69] in the classical context, connecting thus the (free) entropy,Wasserstein distance and Fisher information. As we will see, the HWI implies the Log-Sobolev inequality for strictly convex potentials. This free HWI inequality was not considered before, and in particular it is not clear whether there is a random matrix proof, delicate points involving the Wasserstein distance entering into the proof.

Theorem (6.3.8) [177]: (HWI inequalities)

(i) Assume that V is C^2 such that for some $\rho\in\mathbf{R}$, $V(x)-\rho x^2$ is convex. Then, for any measure $\mu\in \mathrm{P}(\mathbf{R})$

$$\mathrm{E}(\mu)-\mathrm{E}(\mu_V)\leq\sqrt{1(\mu)}\mathrm{W}_2(\mu,\mu_V)-\rho\mathrm{W}_2^2(\mu,\mu_V) \tag{103}$$

In the case $V(x)=\rho x^2$, the inequality is sharp.

(ii) If V is C^2 and $V(x)=\rho x^2$ is convex for some $\rho>0$ and $\mathrm{p}>0$, then for the same constant c_p appearing in Theorem (6.3.2), we have that

$$\mathrm{E}(\mu)-\mathrm{E}(\mu_V)\leq 1_q^{1/q}(\mu)\mathrm{W}_p(\mu,\mu_v)-\rho c_p\mathrm{W}_p^p(\mu,\mu_V) \tag{104}$$

where $1/\mathrm{p}+1/\mathrm{q}=1$.

Proof: (i) We employ here the notations used in Theorem (6.3.7) and we will give a proof of the inequality for the case of a measure μ with smooth and compactly supported density, the general case follows through careful approximations pointed in [66]. The inequality to be proved can be restated as (105)+(106)+ (107) ≥ 0, where

$$\left(\int\left(H\mu(\theta(x))-V'(\theta(x))\right)^2\mu_V(dx)\int\left(\theta(x)-x\right)^2\mu_V(dx)\right)^{1/2}$$

$$-\int \left(H\mu(\theta(x))-V'(\theta(x))\right)(x-\theta(x))\,\mu_V(dx) \tag{105}$$

$$=\int \left[V(x)-V(\theta(x))-V'(\theta(x))(x-\theta(x))-\rho(\theta(x)-x)^2\right]\mu_V(dx) \tag{106}$$

$$=\int H\mu(\theta(x))(x-\theta(x))\,\mu_V(dx)-\int\int \log\frac{x-y}{\theta(x)-\theta(y)}\,\mu_v(dx)\,\mu_V(dy) \tag{107}$$

A simple application of Cauchy's inequality shows that (105) ≥ 0. Using convexity of $V(x)-\rho x^2$ we have from Eq. (101), that (106) ≥ 0. Finally, using (102), we have that

$$\int\int\left(\frac{x-y}{\theta(x)-\theta(y)}-1-\log\frac{x-y}{\theta(x)-\theta(y)}\right)\mu_V(dx)\,\mu_V(dy)\geq 0$$

which finishes the proof of (103). For the case $V(x)=\rho x^2$, we have equality if $\theta(x)=x+m$.

(ii) The inequality we want to prove is equivalent to the statement that (108) + (109) + (110) ≥ 0, where

$$=\left|\int \left|H\mu(\theta(x))-V'(\theta(x))\right|^q \mu_V(dx)\right|^{1/q}\left|\int (\theta(x)-x)^p\,\mu_V(dx)\right|^{1/p}$$

$$-\int \left(H\mu(\theta(x))-V'(\theta(x))\right)(x-\theta(x))\,\mu_V(dx) \tag{108}$$

$$=\int \left[V(x)-V(\theta(x))-V'(\theta(x))(x-\theta(x))-\rho c_p\left|\theta(x)-x\right|^p\right]\mu_V(dx) \tag{109}$$

$$=\int H\mu(\theta(x))(x-\theta(x))\,\mu_V(dx)-\int\int \log\frac{x-y}{\theta(x)-\theta(y)}\,\mu_v(dx)\,\mu_V(dy) \tag{110}$$

Now, (108) is non-negative thanks to Hölder's inequality, Eq. (109), follows from the convexity of $V(x)-\rho|x|^p$ and the combination of (80) and (81), while Eq. (110) is the same as (107).

As pointed out in [69], HWI inequalities for $\rho>0$ always implies Log-Sobolev. We give here the following formal corollary of HWI inequality.

Corollary (6.3.9) [177]:(i) If $\rho > 0$, then inequality (103) implies (98) and (104) implies (97).

(ii) If $V(x)-\rho x^2$ is a convex for some $\rho \in \mathbf{R}$, then Talagrand's free transportation inequality with constant $C > \max\{0,-\rho\}$ implies free Log-Sobolev inequality with constant $K = \max\left\{\rho, \frac{(C+\rho)^2}{32C}\right\}$. More precisely,

$$CW_2^2(\mu,\mu_V) \le E(\mu)-E(\mu_V) \Rightarrow E(\mu)-E(\mu_V) \le \frac{1}{4K}1(\mu), \quad \forall \mu \in P(\mathbf{R})$$

(iii) In particular, if V is convex and C^2 such that $V''(x) \ge \rho > 0$ for $|x| \ge r$, then free Log- Sobolev inequality holds with the constant $C > 0$ from (72).

Proof: (i) It follows as an application of Young's inequality $a^p / p + b^q / q \ge ab$ for $a,b \ge 0$.

(ii) For $\rho > 0$, everything is clear. In the case $\rho \le 0$, then, from (103) and Talagrand's transportation inequality, one has for $\delta > 0$, that

$$\begin{aligned} E(\mu)-E(\mu_V) &\le \sqrt{1(\mu)}W_2(\mu,\mu_V)-\rho W_2^2(\mu,\mu_V) \\ &\le 4\delta I(\mu)+\left(\frac{1}{C\delta}-\frac{\rho}{C}\right)(E(\mu)-E(\mu_V)) \end{aligned}$$

which yields for any $\delta > \frac{1}{c+\rho}$

$$E(\mu)-E(\mu_V) \le \frac{4C\delta^2}{(c+\rho)\delta-1}1(\mu)$$

Taking minimum over $\delta > \frac{1}{c+\rho}$ gives the conclusion.

(iii) In the case V is convex, C^2 and strongly convex for large values, part (ii) of Theorem (6.3.2) does the rest.

The (one-dimensional) free Brunn–Minkowski inequality was put forward in [166] again through random matrix approximation. We provide here a direct mass transportation proof similar to the one of its classical (one-dimensional) counterpart (see [241]). As discussed in [171], this inequality may be used to deduce in an easy way both the Log-Sobolev and transportation inequalities.

We show the following theorem.

Theorem (6.3.10) [177]: Assume that V_1, V_2, V_3 are some potentials satisfying (120) such that for some $a \in (0,1)$

$$aV_1(x)+(1-a)V_2(y) \geq V_3(ax+(1+a)y) \text{ for all } x, y \in \mathbf{R}. \tag{111}$$

Then

$$a\mathrm{E}_{V_1}(\mu_{V_1})+(1-a)\mathrm{E}_{V_2}(\mu_{V_2}) \geq \mathrm{E}_{V_3}(\mu_{V_3}) \tag{112}$$

Proof: Take the (increasing) transportation map θ from μ_{V_1} into μ_{V_2}. This certainly exists as the measure μ_{V_1} has no atoms.

Noticing that for any measure with finite logarithmic energy, we have the obvious equality

$$\int \log|x-y| \mu(dx)\mu(dy) = 2\int_{x>y} \log(x-y)\mu(dx)\mu(dy)$$

Using this we argue that

$$\begin{aligned}
&\int aV_1(x)+(1-a)V_2(\theta(x))\mu_{V_1}(dx) \\
&\qquad -2\int_{x>y}\int \left(a\log(x-y)+(1-a)\log(\theta(x)-\theta(y))\right)\mu_{V_1}(dx)\mu_{V_1}(dy) \\
&\qquad \geq \int V_3(ax+(1+a)\theta(x))\mu_{V_1}(dx) \\
&\qquad -2\int_{x>y}\int \left[\log(ax+(1+a)\theta(x))-(ay+(1-a)\theta(y))\right]\mu_{V_1}(dx)\mu_{V_1}(dy) \\
&\qquad = \mathrm{E}_{V_3}(\mathrm{V}) \geq \mathrm{E}_{V_3}(\mu_{V_3})
\end{aligned}$$

where $\mathrm{V} = (aid+(1-a)\theta)_{\#}\,\mu_{V_1}$ and we used (111) and the concavity of the logarithm on $(0,\infty)$.The proof is complete.

We investigate Poincaré type inequalities in the free (one-dimensional) context. We discuss two versions of it. The first one is suggested by large matrix approximations and the classical Poincaré inequality for strictly convex potentials, but will be proved directly. Recall first the classical Poincaré inequality (cf. e.g. [25, 174, 23, 71]).

Theorem (6.3.11) [177]: Let $\mu(dx) = e^{-W(x)}dx$ be a probability measure on $\mathbf{R}^d$ such that $\mathrm{W}(x)-r|x|^2$ is convex. Then for any compactly supported and smooth function

$\phi : \mathbf{R}^d \to \mathbf{R}$, we have that

$$\int |\nabla\phi|^2 \, d\mu \geq r Var_\mu(\phi) \qquad (113)$$

Assume now that V is a potential on $\mathbf{R}$ with enough growth at infinity. Consider the matrix models on H_n, the space of Hermitian $n \times n$ matrices with the inner product $\langle A, B \rangle = T_r(AB^*)$ and the probability measure given by

$$\mathrm{P}_n(dM) = \frac{1}{\mathrm{Z}_n(V)} e^{-nTr(V(M))} dM$$

where here dM is the standard Lebesgue measure on H_n. We have that for any bounded continuous function $F : \mathbf{R} \to \mathbf{R}$,

$$\int \frac{1}{n} T_r(F(M)) \mathrm{P}_n(dM) \xrightarrow[n\to\infty]{} \int F(x) \mu_V(dx) \qquad (114)$$

Assume in addition that $V(x) - \rho x^2$ is a convex function on $\mathbf{R}$. Then, consider

$\Phi(M) = T_r \phi(M)$, where $\phi : \mathbf{R} \to \mathbf{R}$ is a compactly supported and smooth function. Notice that $\nabla\Phi(M) = \phi'(M)$ and thus

$|\nabla\Phi(M)|^2 = |\phi'(M)|^2 = T_r(\phi'(M)^2)$. Since $nT_r(v(M)) - n\rho|M|^2$ is convex as a function of M, we can apply Poincaré's inequality on H_n to obtain that

$$\int T_r(\phi'(M)^2) \mathrm{P}_n(dM) \geq n\rho Var \mathrm{P}_n(T_r(\phi'(M))) \qquad (115)$$

The first term in this inequality divided by n (cf. Eq. (114)) converges to

$\int \varepsilon'(x)^2 \mu_v(dx)$. To understand the second term in the above equation, notice that $Var(T_r(\phi(M))) = \mathrm{E}\left[\left(T_r(\phi(M)) - \mathrm{E}\left[T_r(\phi(M))\right]\right)^2\right]$. The study of the asymptotic of the linear statistics, $T_r(\phi(M)) - \mathrm{E}\left[T_r(\phi(M))\right]$ in the literature of random matrix is known as "fluctuations". From Johansson's [132], it is known that this is universal in the sense that the limit in distribution of the fluctuations is Gaussian and, at least in the case of polynomial V (for which $V(x) - \rho x^2$ fulfills the conditions in there), the variance of the Gaussian limit depends only on the endpoints of the support of μ_v. Moreover, in the particular case of $V(x) - 2x^2$, the variance of the distribution was computed for example in [11] and [132] as

$$\frac{1}{2\pi^2}\int_{-1}^{1}\int_{-1}^{1}\left(\frac{\phi(t)-\phi(s)}{t-s}\right)^2\frac{1-ts}{\sqrt{1-t^2}\sqrt{1-s^2}}dtds \tag{116}$$

This variance is interpreted in [258] in terms of the number operator of the arcsine law.

Corollary(6.3.12)[274]: $\delta=\frac{(c+\rho)\pm\sqrt{(c+\rho)^2-64\rho c}}{32\rho c}$

Proof: Corollary(6.3.9) and Theorem(6.3.7) show that $\frac{1}{4\rho}I(\mu)\leq\frac{4c\delta^2}{(c+\rho)}I(\mu)$ and

$$16c\,\delta^2-(c+\rho)\delta+1\geq 0$$

$$\delta=\frac{+(c+\rho)\pm\sqrt{(c+\rho)^2-64\rho c}}{32\rho c}$$
$$=\frac{(c+\rho)\pm\sqrt{(c+\rho)^2-64\rho c}}{32\rho c}$$

Corollary(6.3.13)[274]: If V_n, V_{n+1}, V_{n+2} are potentials such that for $0<a<1$

$$aV_n(x)+(1-a)V_{n+1}(x-\varepsilon)\geq V_{n+2}(ax+(1+a)x-\varepsilon)\ \text{forall}\, x\in\mathbf{R}, \varepsilon>0.$$

Then

$$a\mathrm{E}_{V_n}(\mu_{V_n})+(1-a)\mathrm{E}_{V_{n+1}}(\mu_{V_{n+1}})\geq \mathrm{E}_{V_{n+2}}(\mu_{V_{n+2}})$$

Proof: Let $\theta:\mu_{V_n}\to\mu_{V_{n+1}}$ be a transportation map,we have

$$\int\log|\varepsilon|\mu(dx)\mu(dx-\varepsilon)=2\int_{y=x-\varepsilon}\log(\varepsilon)\mu(dx)\mu(dx-\varepsilon)$$

Hence

$$\int aV_n(x)+(1-a)V_{n+1}(\theta(x))\mu_n(dx)$$

$$-2\int_{y=x-\varepsilon}\int\left(a\log(\varepsilon)+(1-a)\log\left(\theta(x)-\theta(x-\varepsilon)\right)\right)\mu_{V_n}(dx)\,\mu_{V_n}(dx-\varepsilon)$$

$$\geq\int V_3\left(ax+(1+a)\theta(x)\right)\mu_{V_1}(dx)$$

$$-2\int_{y=x-\varepsilon}\int\left[\log\left(ax+(1+a)\theta(x)\right)-\left(a(x-\varepsilon)+(1-a)\theta(x-\varepsilon)\right)\right]\mu_{V_n}(dx)\,\mu_{V_n}(dx-\varepsilon)$$

$$=\mathrm{E}_{V_{n+2}}(\mathrm{V})\geq\mathrm{E}_{V_{n+2}}\left(\mu_{V_{n+2}}\right)$$

where $\mathrm{V}=\left(aid+(1-a)\theta\right)_{\#}\mu_{V_n}$ .Theorem (6,3,10)complete the prove

Dividing the inequality in Eq. (115) by n and taking the limit when $\mathrm{n}\to\infty$, these heuristics (after a simple rescaling) suggest the following result.

Theorem (6.3.14) [177]: Assume that $V(x)-\rho x^2$ is convex for some $\rho>0$. Then for any smooth function ϕ, one has that

$$\int\phi'(x)^2\,\mu_V(dx)$$

$$\geq\frac{\rho}{2\pi^2}\iint\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2\frac{-2ab+\left(a+b(x+y)-2xy\right)}{2\sqrt{(x-a)(b-x)}\sqrt{(y-a)(b-y)}}dxdy \qquad (117)$$

where $Supp(\mu_v)=[a,b]$. Equality is attained for $V(x)=\rho(x-x)^2+\beta$ and $\phi(x)=C_1+C_2x$

for some constants C_1 and C_2.

If the numerator in the second fraction of (117) is nonnegative. This is so because

$$-2ab+(a+b)(x-y)-2xy=2\left(\left(\frac{b-a}{2}\right)^2-\left(x-\frac{b-a}{2}\right)\left(y-\frac{b-a}{2}\right)\right)\geq 0$$

for any $x,y,\in[a,b]$.

Proof: Using a simple rescaling we may assume without loss of generality that $\mathrm{a}=-1$ and $\mathrm{b}=1$ and the inequality we have to show reduces to

$$\int \phi'(x)^2 \mu_v(dx) \geq \frac{\rho}{2\pi}\int_{-1}^{1}\int_{-1}^{1}\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2 \frac{1-xy}{\sqrt{1-x^2}\sqrt{1-y^2}}dxdy \tag{118}$$

Then, based on Eq. (124), we have that

$$q(x)=\frac{\sqrt{1-x^2}}{2\pi^2}\int_{-1}^{1}\frac{V'(y)-V'(x)}{\sqrt{1-y^2}(y-x)}dy$$

From the convexity of $V(x)-\rho x^2$, we learn that $\frac{V'(y)-V'(x)}{y-x}\geq 2\rho$ and thus that

$$q(x)\geq\frac{\rho}{\pi}\sqrt{1-x^2} \tag{119}$$

which implies

$$\int \phi'(x)^2 \mu_v(dx)\geq\frac{\rho}{\pi}\int \phi'(x)^2\sqrt{1-x^2}dx$$

Therefore it is enough to check that

$$\int_{-1}^{1}\phi'(x)^2\sqrt{1-x^2}dx \geq \frac{1}{2\pi}\int_{-1}^{1}\int_{-1}^{1}\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2\frac{1-xy}{\sqrt{1-x^2}\sqrt{1-y^2}}dxdy \tag{120}$$

for any smooth ϕ. Now, we make the change of variables $x=\cos t$ to justify

$$\int_{-1}^{1}\phi'(x)^2\sqrt{1-x^2}dx=\int_{0}^{\pi}\phi'(\cos t)^2\sin^2 tdt=\int \psi'(t)^2 dt$$

where $\psi(t)=\phi(\cos t)$.

On the other hand, using the change of variable $x=\cos t \quad y=\cos s$ on the right-hand side, inequality (120) becomes

$$\int_{0}^{\pi}\psi'(t)^2 dt \geq \frac{1}{2\pi}\int_{0}^{\pi}\int_{0}^{\pi}\left(\frac{\psi(t)-\psi(s)}{\cos t-\cos s}\right)(1-\cos t\cos s)dtds \tag{121}$$

To show this, we write $\psi(t)=\sum_{k=0}^{\infty} a_k \cos kt$ and then, because ψ is a smooth function, we can differentiate term by term to get $\psi'(t)=-\sum_{k-1}^{\infty} ka_k \sin kt$, therefore

$$\int_0^\pi \psi'(t)^2\, dt = \frac{\pi}{2}\sum_{k=1}^{\infty} k^2 a_k^2$$

and

$$\int_0^\pi\int_0^\pi \left(\frac{\psi(t)-\psi(s)}{\cos t-\cos s}\right)(1-\cos t\cos s)\, dtds$$

$$=\sum_{k,l=1}^{\infty} a_k a_l \int_0^\pi\int_0^\pi \frac{(\cos kt-\cos ks)(\cos lt-\cos ls)(1-\cos t\cos s)}{(\cos t-\cos s)^2}\, dtds$$

To compute the integrals on the right-hand side of the above equation, we take the generating function of these numbers and with a little algebra one can show that

$$\sum_{k,l=1}^{\infty} u^k v^l \int_0^\pi\int_0^\pi \frac{(\cos kt-\cos ks)(\cos lt-\cos ls)(1-\cos t\cos s)}{(\cos t-\cos s)^2}\, dtds$$
$$=\int_0^\pi\int_0^\pi \frac{(u-u^3)(v-v^3)(1-\cos t\cos s)}{(1+u^2-2u\cos t)(1+u^2-2u\cos s)(1+v^2-2v\cos t)(1+v^2-2v\cos s)}\, dtds \quad (122)$$
$$=\frac{\pi^2 uv}{(1-uv)^2}=\pi^2\sum_{k=1}^{\infty} ku^k v^k$$

for all $u,v\in(-1,1)$. The last integral can be computed as follows. First use partial fractions to justify

$$\int_0^\pi \frac{(A+B\cos t)dt}{(1+u^2-2u\cos t)(1+v^2-2v\cos t)}=\int_0^\pi \frac{Cdt}{1+u^2-2u\cos t}+\int_0^\pi \frac{Ddt}{1+v^2-2v\cos t}=\frac{C/2}{1-u^2}+\frac{D/2}{1-v^2}$$

where the constants C,D are linear combinations of A and B. Further, taking $\mathrm{A}=1$ and $\mathrm{B}=-\cos s$ and repeating once more the partial fractions argument, one can cary out the proof of (122).

The main consequence of the above calculation is that

$$\int_0^\pi\int_0^\pi \frac{(\cos kt-\cos ks)(\cos lt-\cos ls)(1-\cos t\cos s)}{(\cos t-\cos s)^2}\, dtds=\pi^2 k\delta_{kl}$$

and that

$$\int_0^{\pi}\int_0^{\pi}\left(\frac{\psi(t)-\psi(s)}{\cos t-\cos s}\right)(1-\cos t\cos s)\,dtds=\pi^2\sum_{k=1}^{\infty}ka_k^2 \tag{123}$$

Therefore inequality (7.9) becomes equivalent to

$$\frac{\pi}{2}\sum_{k=1}^{\infty}k^2a_k^2\geq\frac{\pi}{2}\sum_{k=1}^{\infty}ka_k^2$$

which is obviously true. Notice that equality in this inequality is attained for the case $a_k=0$ for all $k\geq 2$ and arbitrary a_1. This corresponds to the case $\psi(t)=c_1+c_2\cos t$ or $\phi(x)=c_2x+c_1$ for some c_1,c_2.

Finally we point out that equality in (118) is attained if the equality is attained in (119) and (121). From there one can easily see from rescaling that equality in (117) is attained for $V(x)=\rho(x-x)^2+\beta$ and $\phi(x)=c_1+c_2x$. The proof of Theorem (6.3.14) is complete.

The second version of the Poincaré inequality is motivated by the free calculus and the noncommutative derivative. It was already investigated by Biane [200] for the case of the semicircular law.

Definition (6.3.15) [177]: For a given probability measure μ on $\mathbf{R}$, we say that it satisfies a Poincaré inequality if there is a constant $C>0$ such that

$$\iint\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2\mu(dx)\mu(dy)\geq Var_{\mu}(\phi)\text{ for every }\phi\in C_0^1(\mathbf{R}) \tag{124}$$

By the best constant we mean the largest $C>0$ for which the above inequality is satisfied and we denote it by $\mathrm{P}(\mu)$ or $\lambda_1(\mu)$ or $\mathrm{SG}(\mu)$.

In the noncommutative setting for a given function ϕ, we can think of

$D\phi(x,y)=\dfrac{\phi(x)-\phi(y)}{x-y}$ as the noncommutative derivative of ϕ. As pointed out by Voiculescu in [44], this is the unique $mapD:C\langle x\rangle\to C\langle x\rangle\otimes C\langle x\rangle$ such that

(i) $\mathrm{D}_1=0$.

(ii) $D(fg)=D(f)g+fD(g)$ for any $f,g\in C\langle x\rangle$

First we collect a couple of obvious properties of the Poincaré constant.

Proposition (6.3.16)([177]:(i) For any $a \neq 0$,

$$\mathrm{P}\left((ax+b)_{\#}\,\mu\right)=\frac{1}{a^2}\mathrm{P}(\mu)$$

where here and elsewhere, for a given function $f:\mathbf{R}\to\mathbf{R}$, $f_{\#}\mu$ is the push forward measure given by $(f_{\#}\mu)(A)=\mu\left(f^{-1}(A)\right)$.

(ii) If $f:\mathbf{R}\to\mathbf{R}$ is a differential map such that $\left|f'(x)\right|\geq C>0$ for all $x\in\mathbf{R}$, then

$$\mathrm{P}(\mu)\geq C^2\mathrm{P}(f_{\#}\mu)$$

(iii) If $\{\mu_n\}_{n\geq 1}$ is a sequence of probability measures which converges weakly to μ, then

$$\mathrm{P}(\mu)\geq \limsup_{n\to\infty}\mathrm{P}(\mu_n)$$

Next we describe some bounds for the Poincaré constant.

Theorem (6.3.17) [177]: Assume that the measure μ has compact support and is not concentrated at one point. Then μ satisfies a Poincaré inequality with

$$\frac{2}{d^2(\mu)}\leq \mathrm{P}(\mu)\leq\frac{1}{Var(\mu)} \tag{125}$$

where $d(\mu)=diam\left(\operatorname{supp}(\mu)\right)$ is the diameter of the support of μ and

$$Var(\mu)=\int x^2\mu(dx)-\left(\int x\mu(dx)\right)^2.$$

Equality on the left in (125) is attained only for the case

$$\mu=\alpha\delta_a+(1-\alpha)\delta_b,\qquad a<b,\quad d0<\alpha<1.$$

Equality on the right of (125) is attained only for the case of a semicircular law $(a\in\mathbf{R}, r>0)$

$$\mu(dx)=\frac{1}{2\pi r^2}1_{[a-2r,a+2r]}(x)\sqrt{4r^2-(x-a)^2}\,dx$$

In addition, assume that V is aC^2 potential on $\mathbf{R}$ such that for some integer p and real $\rho>0$, $V(x)-\rho x^{2p}$, is convex and μ is the minimizer of

$$\int V(x)\mu(x)-\iint\log|x-y|\,\mu(dx)\mu(dy)$$

over all probability measures of $\mathbf{R}$. Then

$$\frac{\left(p\rho\binom{2p}{p}\right)^{\frac{1}{p}}}{8} \leq \mathrm{P}(\mu) \tag{126}$$

In particular if $p=1$, we get that $\frac{\rho}{4} \leq \mathrm{P}(\mu)$

Proof: For a given function $\phi \in C_0^1(\mathbf{R})$, the left-hand side of (125) follows from

$$\begin{aligned} Var\mu(\phi) &= \frac{1}{2}\int\int\left(\phi(x)-\phi(y)\right)^2 \mu(dx)\mu(dy) \\ &= \frac{1}{2}\int\int(x-y)^2\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2 \mu(dx)\mu(dy) \\ &\leq \frac{d^2(\mu)}{2}\int\int\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2 \mu(dx)\mu(dy) \end{aligned} \tag{127}$$

The right-hand side of (125) follows from (124) for $a\phi \in C_0^1(\mathbf{R})$ such that $\phi(x)=x$ on the support of μ.

For measures $\mu=\alpha\delta_a+(1-\alpha)\delta_b$, condition (124) is equivalent to

$$C\alpha(1-\alpha)\left(\phi(b)-\phi(a)\right)^2 \leq \alpha^2\left(\phi'(a)\right)^2+(1-\alpha)^2\left(\phi'(b)\right)^2 + 2\alpha(1-\alpha)\left(\frac{\phi(b)-\phi(a)}{b-a}\right)^2 \quad \text{for any } \phi \in C_0^1(\mathbf{R})$$

Since for any function $\phi \in C_0^\infty(\mathbf{R})$ we can find another function $\psi \in C_0^1(\mathbf{R})$ so that $\phi(a)=\psi(a)$ and $\phi(b)=\psi(b)$ and $\psi'(a)=0$, $\psi'(b)=0\psi$, this is also equivalent to

$$C\alpha(1-\alpha)\left(\psi(b)-\psi(a)\right)^2 \leq 2\alpha(1-\alpha)\left(\frac{\psi(b)-\psi(a)}{b-a}\right)^2 \text{ for any } \psi \in C_0^1(\mathbf{R})$$

This amounts to $C \leq 2/(b-a)^2$ and therefore, in this case, $Poin(\mu)=\frac{2}{d^2(\mu)}$.

Conversely, if μ is a measure so that $\mathrm{P}(\mu)=\frac{2}{d^2(\mu)}$, then, for $1>\varepsilon>0$, there is a function $\phi_\varepsilon \in C_0^1(\mathbf{R})$ such that

$$\left(\frac{2}{d^2(\mu)}+\varepsilon^2\right)Var_\mu(\phi_\varepsilon) > \iint\left(\frac{\phi_\varepsilon(x)-\phi_\varepsilon(y)}{x-y}\right)\mu(dx)\mu(dy)$$

Without loss of generality we can assume that $0=\inf\operatorname{supp}(\mu)$, $1=\sup\operatorname{supp}(\mu)$ and $\int\phi_\varepsilon d\mu=0$, $\int\phi_\varepsilon^2 d\mu=1$ where we recall that $\operatorname{supp}(\mu)$ stands for the support of μ. In this case, the above inequality implies

$$2+\varepsilon^2 \geq \iint_{|x-y|\geq 1-\varepsilon}\left(\frac{\phi_\varepsilon(x)-\phi_\varepsilon(y)}{x-y}\right)^2\mu(dx)\mu(dy)+\iint_{|x-y|\square 1-\varepsilon}\left(\frac{\phi_\varepsilon(x)-\phi_\varepsilon(y)}{x-y}\right)^2\mu(dx)\mu(dy)$$

$$\geq \iint_{|x-y|\geq 1-\varepsilon}(\phi_\varepsilon(x)-\phi_\varepsilon(y))^2\mu(dx)\mu(dy)+\frac{1}{(1-\varepsilon)^2}\iint_{|x-y|\square 1-\varepsilon}(\phi_\varepsilon(x)-\phi_\varepsilon(y))^2\mu(dx)\mu(dy)$$

$$=\frac{\varepsilon(2-\varepsilon)}{(1-\varepsilon)^2}\iint_{|x-y|\square 1-\varepsilon}(\phi_\varepsilon(x)-\phi_\varepsilon(y))^2\mu(dx)\mu(dy)+\frac{2}{(1-\varepsilon)^2}$$

which results with

$$\iint_{|x-y|\geq 1-\varepsilon}(\phi_\varepsilon(x)-\phi_\varepsilon(y))^2\mu(dx)\mu(dy)\geq 2-\frac{\varepsilon(1-\varepsilon)}{1-\varepsilon} \tag{128}$$

Now,

$$\iint_{|x-y|\geq 1-\varepsilon}(\phi_\varepsilon(x)-\phi_\varepsilon(y))^2\mu(dx)\mu(dy)\leq \iint_{\substack{|x-1/2|\geq 1/2-\varepsilon\\ |y-1/2|\geq 1/2-\varepsilon}}(\phi_\varepsilon(x)-\phi_\varepsilon(y))^2\mu(dx)\mu(dy)$$

$$\leq 2\mu(|x-1/2|\geq 1/2-\varepsilon) \tag{129}$$

Thus (128) and (129) give $\mu(|x-1/2|\geq 1/2-\varepsilon)\geq 1-\frac{\varepsilon(2-\varepsilon)}{4-2\varepsilon}$ for any $1>\varepsilon>0$.

This shows that $\mu((0,1))=0$ and therefore $\mu=\alpha\delta 0+(1-\alpha)\delta_1$.

The other extreme case of inequality (125) is contained in Biane's [200] in the more general context of several noncommutative variables. For completeness we will provide here a self contained proof. In the first place, we may assume that

$$\mu(dx)=\frac{1}{2\pi}\mathrm{I}_{[-2,2]}x\sqrt{4-x^2dx}$$

is the semicircular law on [−2, 2]. Take U_n to be the Chebyshev polynomials of second kind defined by $U_n(\cos(\theta)) = \frac{\sin(n+1)\theta}{\sin\theta}$. With this choice, we have that $U_n\left(\frac{x}{2}\right)$ are the orthogonal polynomials with respect to μ. The generating function of U_n is given by

$$\sum_{n=0}^{\infty} r^n U_n(x) = \frac{1}{1-2rx+r^2}, \text{ for } |x|,|r|<1$$

from which one gets

$$\sum_{n=0}^{\infty} r \frac{U_n(x)-U_n(y)}{x-y} = \frac{2r}{(1-2rx+r^2)(1-2ry+r^2)} = \sum_{n=0}^{\infty} r^n \sum_{k=0}^{n-1} U_k(x)U_{n-1-k}(y)$$

and then

$$\frac{U_n(x)-U_n(y)}{x-y} = 2\sum_{k=0}^{n-1} U_k(x)U_{n-1-k}(y) \tag{130}$$

Now, for a given $\phi \in C_0^1(\mathbf{R})$, we can write in $L^2(\mu)$ sense,

$$\phi(x) = \sum_{n=0}^{\infty} \alpha_n U_n\left(\frac{x}{2}\right),$$

yielding from orthogonality and (130) that

$$Var\mu(\phi) = \int \phi^2 d\mu - \left(\int \phi d\mu\right)^2 = \sum_{n=1}^{\infty} \alpha_n^2 \text{ and}$$

$$\iint \left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2 \mu(dx)\mu(dy) = \sum_{n=1}^{\infty} n\alpha_n^2$$

It follows that in this case $\mathrm{P}(\mu) = 1 = 1/Var(\mu)$ and equality is attained only for $\phi(x) = c_1 + c_2 U_1(x) = c_1 + c_2 x$ for some constants c_1, c_2.

To prove the converse, take a compactly supported measure μ and assume that $\int x\mu(dx) = 0$ and $\int x^2 \mu(dx) = 1$. In order to show that μ is the semicircular distribution, it suffices to show that $\int U_n\left(\frac{x}{2}\right)\mu(dx) = 0$ for all $n \geq 1$. We use induction to this task. Assuming true for $U_1, U_{,2} ..., U_n$, and using

$U_n(x)2xU_n(x)-U_{n-1}(x)$, we need to show that $xU_n(x/2)$ integrates to 0 against μ. Applying Poincaré's inequality to $U_n(x/2)+eU_1(x/2)$ together with the induction hypothesis and equation (130), we get that for any $r\in\mathbf{R}$,

$$\int U_n^2(x/2)\mu(dx)+r\int xU_n(x/2)\mu(dx)\geq\int\int\left(\frac{U_n(x/2)-U_n(y/2)}{x-y}\right)^2\mu(dx)\mu(dy)$$

which implies that $\int xU_n(x/2)\mu(dx)=0$

In the case of the equilibrium measure of a convex potential V, we have the support of the measure consists of one interval $[\mathrm{a},\mathrm{b}]$ and a,b solve the system (cf. (66))

$$\frac{1}{2\pi}\int_a^b V'(x)\sqrt{\frac{x-a}{b-x}}dx=1 \text{ and } \frac{1}{2\pi}\int_a^b V'(x)\sqrt{\frac{b-x}{x-a}}dx=-1$$

If we denote $\mathrm{c}=(\mathrm{b}-\mathrm{a})/2$ and $\beta=(\mathrm{a}+\mathrm{b})/2$, the system above can be rewritten in terms of β and c as

$$\frac{c}{2\pi}\int_{-1}^1 V'(\beta+ct)\frac{1+t}{\sqrt{1-t^2}}dt=1 \text{ and } \frac{c}{2\pi}\int_{-1}^1 V'(\beta+ct)\frac{1-t}{\sqrt{1-t^2}}dt=-1$$

which is equivalent to

$$\frac{c}{2\pi}\int_{-1}^1 V'(\beta+ct)\frac{t}{\sqrt{1-t^2}}dt=1 \text{ and } \int_{-1}^1 V'(\beta+ct)\frac{1}{\sqrt{1-t^2}}dt=0$$

Since V is C^2 the first equation can be integrated by parts to get that

$$\frac{c^2}{2\pi}\int_{-1}^1 V''(\beta+ct)\sqrt{1-t^2}dt=1$$

On the other hand we know that $V''(x)\geq 2p(2p-1)\rho x^{2p-2}$, hence

$$1 \geq \frac{2p(2p-1)\rho c^2}{2\pi} \int_{-1}^{1} (ct+\beta)^{2p-2} \sqrt{1-t^2}dt$$

$$\geq \frac{2p(2p-1)\rho c^{2p}}{2\pi} \int_{-1}^{1} t^{2p-2} \sqrt{1-t^2}d = \frac{p(2p-1)\rho c^{2p} \binom{2p}{p}}{4p(2p-1)} t = \frac{p\rho \binom{2p}{p} c^{2p}}{4p}$$

This yields

$$c \leq 2\left(m\rho \binom{2p}{p}\right)^{-\frac{1}{2p}}$$

Finally, because $d(\mu) = b - a = 2c$, we arrive at (126).

To conclude this section, we present an inequality which relates the equilibrium measure of a strong convex potential and the arcsine law.

Theorem (6.3.18) [177]: Assume that $V(x) - \rho x^2$ is a convex for some $\rho > 0$ and the equilibrium measure μ_v has support $[a, b]$. Let

$\text{arcsin}\, e_{a,b} = \mathrm{I}_{[a,b]}(x) \dfrac{1}{\pi\sqrt{(b-x)(x-a)}} dx$ be the arcsine law with support $[a, b]$.

Then for any smooth function supported on $[a, b]$,

$$\int \phi'(x)^2 \mu_v(dx) \geq \rho Var_{\text{arcsin}\, e_{a,b}}(\phi) \tag{131}$$

where the variance is considered with respect to the $\text{arcsine}_{a,b}$ law.

Proof: It suffices to deal with the case $a = -1,\ b = 1$, the rest following by simple rescaling. Recall that in the proof of Theorem (6.3.14) , we use convexity to get that the density $g(x)$ of μ_v satisfies $g(x) \geq \frac{\rho}{\pi}\sqrt{1-x^2}$. Thus the proof reduces to

$$\frac{1}{\pi}\int_{-1}^{1} \phi'(x)^2 \sqrt{1-x^2}\,(dx) \geq Var_{\text{arcsin}\, e}(\phi) \tag{132}$$

For this, write $\phi = \sum_{n=0}^{\infty} \alpha_n T_n(x)$ the expansion of ϕ in terms of Chebyshev polynomials of the first kind. Now, $T_n = nU_{n-1}$ and thus the above inequality reduces to the obvious inequality $\sum_{n=1}^{\infty} n^2\alpha_n^2 \geq \sum_{n=1}^{\infty} \alpha_n^2$

We will actually see below that inequality (132) is simply the spectral gap for the Jacobi operator associated to the arcsine law.

We show how the two versions of the Poincaré inequalities can be viewed as spectral gaps for some Jacobi operators. This discussion is mainly driven from the work [258] by Cabanal-Duvillard and his interpretation of the variance in (127) in terms of the number operator of the Jacobi operator associated to the arcsine law. This viewpoint allows for an unified perspective of the Poincaré inequalities presented in the preceding sections.

For our purpose we consider here the Jacobi operators given, for smooth functions on $(-1,\ 1)$,by

$$L_\lambda f(x) = -\left(1-x^2\right) f''(x) + (2\lambda+1) x f'(x) \tag{133}$$

for $\lambda \geq 0$. We consider the Gegenbauer polynomials $C_n^\lambda, \lambda > 0$, defined by the generating function

$$\sum_{n=0}^{\infty} r^n C_n^\lambda(x) = \frac{1}{\left(1-rx+r^2\right)^\lambda}$$

For $\lambda = 0$ we set $C_n^\lambda(x) = T_n(x)/n, n \geq 1$, where T_n are the Chebyshev polynomials of the first kind.

It is known that C_n^λ are eigenfunctions of L_λ, with eigenvalue $n(n+2\lambda)$, i.e.

$$L_\lambda C_n^\lambda = n(n+2\lambda) C_n^\lambda$$

On the other hand the Gegenbauer polynomials are orthogonal with respect to the probability measure

$$\mathrm{V}_\lambda = \frac{2^{2\lambda}\Gamma^2(\lambda+1)}{\pi\,\Gamma(2\lambda+1)} \mathrm{I}_{[-1,1]}(x)\left(1-x^2\right)^{\lambda-1/2}$$

Notice that in the case of $\lambda = 0$, this becomes the arcsine law and for $\lambda = 1$, this is the semicircular law, while for $\lambda = 1/2$, this becomes the uniform measure on $[-1,\ 1]$.

Take now the normalized Gegenbauer polynomials $\phi_n^\lambda = G_n^\lambda / \sqrt{c_n^\lambda}$, where

$c_n^\lambda = \int G_n^\lambda(x)^2 \mathrm{V}_\lambda(dx)$. Then ϕ_n^λ form an orthonormal basis of $L^2(\mathrm{V}_\lambda)$ and thus the operator L_λ is diagonalized in this basis. Consider N_λ to be the counting number operator with respect to the basis ϕ_n^λ, i.e.

$$N_\lambda \phi_n^\lambda = n\phi_n^\lambda \tag{134}$$

This implies that $L_\lambda = N_n^\lambda + 2\lambda N_\lambda$. Therefore we have the following two inequalities

$$L_\lambda \geq (2\lambda+1) N_\lambda \text{ and } N_\lambda \geq 1 - \mathrm{P}_\lambda \tag{135}$$

where P_λ here stands for the projection on constant functions in $L^2(\mathrm{V}_\lambda)$. In other words, $\mathrm{P}_\lambda \phi = \int \phi \mathrm{V}_\lambda$.

Notice that Eq. (135) include two statements. The first one is the comparison of L and N, with the spectral gap $2\lambda+1$ while the second one is the spectral gap of the counting number operator with the spectral gap 1. In the sequel we want to translate these spectral gaps in terms of Poincaré type inequality. For this matter we need to find the kernel of the operator N .Then we have for any function in the domain of definition of L_λ, that $\phi = \sum_{n=0}^{\infty} \alpha_n \phi_n^\lambda$, and then

$$(L\phi, \phi)_{L^2(V_\lambda)} = \sum_{n=0}^{\infty} n(n+2\lambda)\alpha_n^2$$

On the other hand, using integration by parts, we can justify that

$$(L\phi, \phi)_{L^2(V_\lambda)} = \int \phi L_\lambda \phi dv_\lambda = \int \phi'(x)^2 (1-x^2) v_\lambda(dx).$$

For the number operator, we have that

$$\int \phi L_\lambda \phi dv_\lambda = \sum_{n=0}^{\infty} n\alpha_n^2 = \lim_{r\uparrow 1} \sum_{n=0}^{\infty} nr^{n-1}\alpha_n^2$$

Now, for $-1 < \mathrm{r} < 1$,

$$\sum_{n=1}^{\infty} nr^{n-1}\alpha_n^2 = \int\int \phi(x)\phi(y) \sum_{n=1}^{\infty} nr^{n-1} \phi_n^\lambda(x) \phi_n^\lambda(y) \mathrm{V}_\lambda(dx)\mathrm{V}_\lambda(dy).$$

Furthermore, since $\int \phi_n^\lambda d\mathrm{V}_\lambda = 0$ for $n \geq 1$, we also obtain that

$\sum_{n=0}^{\infty} nr^{n-1}\alpha_n^2 = \int\int \phi^2(x)\phi_n^\lambda(y)\mathrm{V}_\lambda(dx)\mathrm{V}_\lambda(dy) = 0$ for $n \geq 0$ and thus, denoting

$$K_\lambda(r,x,y) = -\sum_{n=0}^{\infty} nr^{n-1}\phi_n^\lambda(x)\phi_n^\lambda(y).$$

$$\int\int \phi(x)\phi(y)\sum_{n=0}^{\infty} nr^{n-1}\phi_n^\lambda(x)\phi_n^\lambda(y)\mathrm{V}_\lambda(dx)\mathrm{V}_\lambda(dy) = \frac{1}{2}\int\int(\phi(x)-\phi(y))^2 K_\lambda(r,x,y)\mathrm{V}_\lambda(dx)v_\lambda(dy)$$

The following formula is essentially due to Watson [91] and valid for λ > 0,

$$\sum_{n=1}^{\infty} r^n\phi_n^\lambda(x)\phi_n^\lambda(y) = \frac{(1-r^2)\Gamma(2\lambda)}{2^{2\lambda-1}\Gamma^2(\lambda)}\int_{-1}^{1}\frac{(1-r^2)^{\lambda-1}}{\left(1-2r\left(xy+z\sqrt{(1-x^2)(1-y^2)}\right)+r^2\right)^{1+\lambda}}dz\,.$$

For $\lambda = 0$, we have to deal with the Chebyshev polynomials of the first kind which was more or less what appeared in the proof of Theorem (6.3.14). For this case, we have that (denoting x = cos t and y = cos s),

$$\sum_{n=0}^{\infty}\frac{r^n}{c_n}T_n(x)T_n(y) = \frac{1-r\cos(t+s)}{1-2r\cos(t+s)+r^2}+\frac{1-r\cos(t-s)}{1-2r\cos(t-s)+r^2}\,.$$

where $c_n = \int T_n^2 dv_0 = 1$ for n = 0 and 1/2 otherwise.

Thus, we obtain, after differentiation with respect to r and then limit over $r\uparrow 1$, that

$$K_\lambda(x,y) = \lim_{r\uparrow 1} K_\lambda(r,x,y) = \begin{cases} \dfrac{r(2\lambda)}{2^{3\lambda-1}\Gamma^2(\lambda)}\displaystyle\int_{-1}^{1}\frac{(1-z^2)^{\lambda-1}}{\left(1-xy-z\sqrt{(1-x^2)(1-y^2)}\right)^{1+\lambda}}dz, & \lambda > 0 \\ \dfrac{1-xy}{(x-y)^2}, & \lambda = 0 \\ \dfrac{1}{2(x-y)^2}, & \lambda = 1 \end{cases} \qquad (136)$$

The integrand is not a rational function. In some cases, it is algebraic since $\lambda \geq 0$ need not be an integer.

To reveal the singularity of this kernel, we make the change of variable

$$1-xy-yz\sqrt{(1-x^2)(1-y^2)} = t(1-xy)-\sqrt{(1-x^2)(1-y^2)}\,.$$

Then, after simple algebraic manipulations, setting $f_\lambda:(0,1)\to\mathbf{R}$,

$$f_\lambda(u) = \int_1^{1/u}\frac{\left[(t-1)(1-ut)\right]^{\lambda-1}}{t^{\lambda+1}}dt\,.$$

and

$$H_\lambda(x,y)=\begin{cases}\dfrac{r(2\lambda)\left(1-xy+\sqrt{(1-x^2)(1-y^2)}\right)^{\lambda}}{2^{3\lambda-1}\Gamma^2(\lambda)\left((1-x^2)(1-y^2)\right)^{\lambda-1/2}}f_\lambda\left(\dfrac{(x-y)^2}{\left(1-xy+\sqrt{(1-x^2)(1-y^2)}\right)}\right), & \lambda>0\\ 1-xy, & \lambda=0\\ \dfrac{1}{2}, & \lambda=1\end{cases} \quad (137)$$

We can rewrite Eq (134) for $|x|,|y|<1$ as

$$K_\lambda(x,y)=\frac{H_\lambda(x,y)}{(x-y)^2}. \quad (138)$$

where $K_\lambda(x,y)$ is a continuous function of $x,y\in[-1,1]$.

Corollary(6.3.19)[274]: $\|\mathbf{P}_0\|\le\int\frac{\Gamma(1)}{\pi}I_{[-1,1]}(x)(1-x^2)^{-\frac{1}{2}}$ Since

$\mathbf{P}_\lambda\,\phi=\int\phi\mathrm{V}_\lambda$,and $|\mathbf{P}_\lambda\,\phi|=\int|\phi||\mathrm{V}_\lambda|$,

then

$$\|\mathbf{P}_\lambda\|\le\sup\int|\phi||\mathrm{V}_\lambda|=\int|\mathrm{V}_\lambda|=\int\frac{\Gamma^2(\lambda+1)}{\pi\Gamma(\lambda+1)}I_{[-1,1]}(x)(1-x^2)^{\lambda-\frac{1}{2}}.$$

Hence

$$\|\mathbf{P}_0\|\le\int\frac{\Gamma(1)}{\pi}I_{[-1,1]}(x)(1-x^2)^{-\frac{1}{2}}.$$

We can find the norm of the projection of $\mathbf{P}_\lambda$ of $\lambda=\frac{1}{2},1$ in the interval $[-1,1]$.

Now, from (136), we obtain the following result.

Theorem (6.3.20) [177]: For any $\lambda\ge0$, one has for all $\lambda\ge0$ and any $\phi\in([-1,1])$, that

$$\int\phi'(x)^2(1-x^2)\mathrm{V}_\lambda(dx)\ge\frac{2\lambda+1}{2}\iint\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2H_\lambda(x,y)\mathrm{V}_\lambda(dx)\mathrm{V}_\lambda(dy). \quad (139)$$

And

$$\iint\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2 H_\lambda(x,y)\mathrm{V}_\lambda(dx)\mathrm{V}_\lambda(dy)\geq 2Var_{\mathrm{V}_\lambda}(\phi) \qquad (140)$$

Remark (6.3.21) [177]:

Combining Eqs. (139) and (140), we also get a Brascamp–Lieb type inequality:

$$\int\phi'(x)^2\left(1+x^2\right)\mathrm{V}_\lambda(dx)\geq(2\lambda+1)Vra_{\mathrm{V}_\lambda}(\phi). \qquad (141)$$

For $\lambda\geq 1/2$, the measure V_2 is of the form $e^{-V(x)}dx$, where

$V(x)=-c_\lambda-(\lambda-1/2)\times\log\left(1-x^2\right)$, a strictly convex function on $(-1,\ 1)$ and according to the classical Brascamp–Lieb inequality [286],

$$\int\phi'(x)^2\frac{\left(1-x^2\right)}{\left(1+x^2\right)}\mathrm{V}_\lambda(dx)\geq(2\lambda-1)Vra_{\mathrm{V}_\lambda}(\phi)\ . \qquad (142)$$

Notice here that neither (141) not (142) implies the other which means that they complement each other in some sense. For example if ϕ has support in $\left[-\frac{1}{2\lambda},\frac{1}{2\lambda}\right]$ (141) implies (142) while if ϕ is supported on $[-1,1]\Big\backslash\left[-\frac{1}{2\lambda},\frac{1}{2\lambda}\right]$, (142) implies (141),

We address the preceding functional inequalities for probability measures on the real positive axis in the context of the Wishart Ensembles from random matrix theory and their associated Marcenko–Pastur distributions.

We start with the random matrix heuristics although, as far as we know, it has not been used towards functional inequalities as before. The problems of large deviations principle for the distribution of the eigenvalues of Wishart ensembles is discussed in [64]. The model is as follows.

Take $T(n)$ a $n\times p(n)$ random matrix with all the entries being iid $N(0,\ 1)$ random variables.

Then $T(n)T(n)^t$ for $n<p(n)$ is known as the nonsingular Wishart random ensemble. According to [65], the distribution of the Wishart ensembles is given by

$$C_{np}e^{-\frac{p(n)}{2}T_rM}\left(\det M\right)^{(p-n)/2}dM\ .$$

where the measure $dM = \prod_{i \le j} dM_{ij}$ the restriction of the Lebegue measure on the set of $n \times n$ non-negative matrices.

It is also known (for example [65]) that the joint distribution of eigenvalues $(\lambda_1, \lambda_2, ..., \lambda_n)$ of $\frac{1}{p(n)} T(n) T(n)^t$ is given by

$$\frac{1}{Z_n} e^{-\frac{p(n)}{2} \sum_{i=1}^n t_i \prod_{i=1}^n \lambda_i^{(p(n)-n-1)/2} \prod_{1 \le i \le j \le n} |\lambda_i - \lambda_j|}.$$

Our interest is in the limit distribution of $\mu_n = \frac{1}{n} \sum_{i=1}^n \lambda_i$. The classical result states that if $n / p(n) \to \alpha \in (0,1]$, then the limit distribution of μ_n is the so called Marcenko–Pastur distribution given by

$$\mathbf{1}_{\left[(1-\sqrt{\alpha})^2, (1+\sqrt{\alpha})^2\right]}(x) \frac{\sqrt{4x - (x - 1x)^2}}{2\pi\alpha x} dx .$$

This is a particular model for the standardWishart ensembles. However one can consider a more general example with potentials for which the distribution of the matrix is driven by a potential $Q : [0, \infty) \to \mathbf{R}$,

$$C_n e^{-p(n) T_r Q(M)} (\det M)^{\lambda(n)} dM .$$

where dM stands for the Lebesgue measure on $\mathrm{n} \times \mathrm{n}$ positive definite matrices. The distribution of eigenvalues of M is given by

$$\frac{1}{Z_n} e^{-p(n) \sum_{i=1}^n Q(t_i) \prod_{i=1}^n t_i^{\lambda(n)} \prod_{1 \le i \le j \le n} |t_i - t_j|}.$$

The main result of [64] is that the distribution of the random measures $\mu_n = \frac{1}{p(n)} \sum_{i=1}^{p(n)} \delta_{\lambda_i}$ under the conditions

$n / p(n) \xrightarrow[n \to \infty]{} \alpha \in (0,1], \lambda(n) / n \xrightarrow[n \to \infty]{} \lambda > 0, v_n$ satisfy a large deviation principle with scale n^{-2} and the rate function given by

$$R(\mu) = \tilde{\mathrm{E}}_Q(\mu) - \inf_{\mu \in \mathrm{P}((0,\infty))} \tilde{\mathrm{E}}_Q(\mu).$$

Where

$$\tilde{E}_Q(\mu) = \alpha \int (Q(x) - \gamma \log(x)) \mu(dx) - \frac{\alpha^2}{2} \int \int \log|x-y| \mu(dx) \mu(dy).$$

This gives the following motivation. Assume that $V : [0,\infty) \to \mathbf{R} \cup \{+\infty\}$ is a lower semicontinuous potential such that $\lim_{|x|\to\infty} (V(x) - 2\log|x|) = \infty$. Then, according to the results in [52], we know that there is a unique minimizer of

$$\inf_{\mu \in P((0,\infty))} E_V(\mu).$$

In addition the equilibrium measure μ_v has compact support.

A particular case of interest is $V(x) = rx - \log(x)$ with $r > 0, \quad s \geq 0$ for which we know [52] that the equilibrium measure is given by

$\mu_V(dx) = 1_{[a,b]}(x) \frac{r\sqrt{(x-a)(b-x)}}{2\pi x} dx$ where

$$a = \frac{s+2-2\sqrt{s+1}}{r}, \qquad b = \frac{s+2+2\sqrt{s+1}}{r}. \qquad (143)$$

One recovers the Marcenko–Pastur distribution for $V(x) = rx - \log(x)$, $r > 0, \quad s \geq 0$, with $r = 1/\alpha$ and $s = (1-\alpha)/\alpha$.

The natural way to deal with functional inequalities in the context of measures on the positive axis [0,∞) is to transfer measures from [0,∞) into measures on the whole $\mathbf{R}$. For a measure μ on [0,∞), consider thus the associated symmetric measure $\tilde{\mu}$ on $\mathbf{R}$ defined as

$$\mu(F) = \tilde{\mu}(\{x : x^2 \in F\}). \qquad (144)$$

for any measurable set F of [0,∞). Defining $\tilde{V}(x) = V(x^2)/2$, it is then an easy exercise to check that

$$E_v(\mu) = 2E_{\tilde{v}}(\tilde{\mu}). \qquad (145)$$

In addition, the minimizer of $E_{\tilde{V}}$ is $\mu_{\tilde{V}} = \tilde{\mu}_V$. Further, for the non-decreasing transportation map θ of μ_V into μ, define

$$\tilde{\theta}(x) = sign(x)\sqrt{\theta(x^2)} \qquad (146)$$

which transports $\tilde{\mu}_{\tilde{V}}$ into $\tilde{\mu}$.

In addition, as it was pointed out in [66], the relative free Fisher information $I_V(\mu)$ is defined for measures μ on $[0,\infty)$ with density $p = d\mu dx$ in $L^3([0,\infty), xdx)$ as

$$I_V(\mu) = \int_0^\infty x\left(H_\mu(x) - V'(x)\right)^2 \mu(dx) \text{ with } H_\mu(x) = p.v.\int \frac{2}{x-y}\mu(dy) \quad (147)$$

Otherwise we take $I_V(\mu) = +\infty$. The main reason for defining this in this way is because, cf. [142] and the discussion following, one has

$$I_V(\mu) = 2I_{\tilde{V}}(\tilde{\mu}) \quad (148)$$

where $I_{\tilde{V}}$ is defined by (154).

To state the transportation cost result, we define the appropriate distance. For any $\mu, \mathrm{V} \in \mathrm{P}([0,\infty))$, set the distance as

$$\mathrm{W}(\mu, \mathrm{V}) = \inf_{\pi \in \prod(\mu,\nu)} \left(\int \left(\sqrt{x} - \sqrt{y}\right)^2 \pi(dx, dy) \right) \quad (149)$$

where $\Pi(\mu, \mathrm{V})$ is the set of probability measures on $\mathbf{R}^2$ with marginals μ and ν.

We have the following transportation cost inequality.

Theorem (6.3.22) [177]: Assume that $V:(0,\infty) \to \mathbf{R}$ is $C^2((0,\infty))$((0, ∞)) such that $V(x^2) - \rho x^2$ is convex on (0,∞) for some $\rho > 0$ and let μ_v be the equilibrium measure of V on [0,∞). Then, for any probability measure μ on [0,∞), we have that

$$\rho \mathrm{W}^2(\mu, \mu_V) \le \mathrm{E}_V(\mu) - \mathrm{E}_V(\mu_V) \quad (150)$$

In the case of $V(x) = rx - s\log(x)$ with r > 0 and $s \ge 0$, this inequality with $\rho = r$ is sharp.

Proof: As announced, the idea is to interpret this inequality as an inequality for potentials on the whole real line instead of $[0,\infty)$. Using the measures $\tilde{\mu}$ and $\tilde{\mu}_v$ from Eq. (144) together with (145), we have that

$$\mathrm{E}_V(\mu) - \mathrm{E}_V(\mu_V) = 2\left(\mathrm{E}_{\tilde{V}}(\tilde{\mu}) - \mathrm{E}_{\tilde{V}}(\tilde{\mu}_V)\right)$$

On the other hand, if θ is the (increasing) transportation map of μ_v into μ, then it is not hard to check that

$$\mathrm{W}^2(\mu,v)=\int\left(\sqrt{x}-\sqrt{\theta(x)}\right)^2\mu_V(dx)=\int\left(x-\tilde{\theta}(x)\right)^2\tilde{\mu}_V(dx)$$

In this framework the inequality (150) translates as

$$\rho\mathrm{W}_2^2(\tilde{\mu},\tilde{\mu}_v)\le E_{\tilde{v}}(\tilde{\mu})-E_{\tilde{v}}(\tilde{\mu}_v) \tag{151}$$

From here we will use the same argument as in the proof of Theorem (6.3.20), start with

$$\begin{aligned}\mathrm{E}_{\tilde{v}}(\tilde{\mu})-\mathrm{E}_{\tilde{v}}(\tilde{\mu}_v)&=\int\left(\tilde{V}\left(\tilde{\theta}(x)\right)-\tilde{V}(x)-\tilde{V}'(x)\left(\tilde{\theta}(x)-x\right)\right)\tilde{\mu}_v(dx)\\&+\iint\left(\frac{\tilde{\theta}(x)-\tilde{\theta}(y)}{x-y}-1-\log\frac{\tilde{\theta}(x)-\tilde{\theta}(y)}{x-y}\right)\tilde{\mu}_v(dx)\tilde{\mu}_v(dy)\end{aligned}$$

and notice that the second line of this is non-negative. For the first line we point out that because $\tilde{V}(x)-\frac{\rho}{2}x^2$ is convex and x and $\tilde{\theta}(x)$ have the same sign, for any X,

$$\tilde{V}\left(\tilde{\theta}(x)\right)-\tilde{V}(x)-\tilde{V}'(x)\left(\tilde{\theta}(x)-x\right)\ge\frac{\rho}{2}\left(\tilde{\theta}(x)-x\right)^2.$$

which implies (150).

In the case $V(x)=rx-s\log(x)$, take $\theta(x)=\left(\sqrt{x}+m\right)^2$ for large m and notice that $\tilde{\theta}(x)=x+msign(x)$. Therefore inequality (10.9) becomes

$$\begin{aligned}rm^2\le rm^2+2rm\int|x|\tilde{\mu}(dx)-2s\int\log\left(\frac{|x+msign(x)|}{|x|}\right)\tilde{\mu}(dx)\\-\iint\log\left(1+m\frac{sign(x)-sign(y)}{x-y}\right)\tilde{\mu}(dx)\tilde{\mu}(dy)\end{aligned}$$

which is sharp for large m.

The next result is the Log-Sobolev type inequality, which was conjectured by Cabanal-Duvillard in [259] for the case of Marcenko–Pastur distribution.

Theorem (6.3.23) [177]: Let V be as in the previous theorem. Then, with the definition from (147) and for any measure $\mu\in\mathrm{P}([0,\infty))$,

$$\mathrm{E}_V(\mu)-\mathrm{E}_V(\mu_V)\le\frac{1}{2\rho}I_V(\mu) \tag{152}$$

In the case $V(x)=rx-s\log(x)$, r >0 and $s\geq 0$ inequality (152) with ρ = r is sharp.

Proof: We will discuss here the proof only in the case when μ has a smooth compactly supported density, careful approximations being described in [66].

From (148), we have $I_v(\mu)=2I_{\tilde{v}}(\tilde{\mu})$, where $I_{\tilde{v}}(\tilde{\mu})=\int\left(H\tilde{\mu}(x)-\tilde{V}'(x)\right)^2\mu(dx)$. Rewriting everything in terms of $\tilde{\mu}$ and the associated quantities, the inequality to be proven can be written in the same way as we did in the proof of Theorem (6.3.8).

$$\begin{aligned}
&\frac{1}{2p}\left(\int H\tilde{\mu}(\theta(x))-\tilde{V}'(\tilde{\theta}(x))\right)^2\tilde{\mu}_v(dx)\\
&+\int\left(\tilde{V}(x)-\tilde{V}'(\tilde{\theta}(x))-\tilde{V}'(\tilde{\theta}(x))(x-\tilde{\theta}(x))\right)\tilde{\mu}_{\tilde{v}}(dx)\\
&-\int\left(H\tilde{\mu}(\tilde{\theta}(x))-\tilde{V}'(\tilde{\theta}(x))\right)(x-\tilde{\theta}(x))\tilde{\mu}_{\tilde{v}}(dx)\\
&+\int H\tilde{\mu}(\tilde{\theta}(x))(x-\tilde{\theta}(x))\tilde{\mu}_{\tilde{v}}(dx)-\int\int\log\frac{x-y}{\tilde{\theta}(x)-\tilde{\theta}(y)}\tilde{\mu}_{\tilde{v}}(dx)\tilde{\mu}_{\tilde{v}}(dy)\geq 0
\end{aligned}\tag{153}$$

Notice that $\tilde{V}(x)-\frac{\rho}{2}x^2$ is not convex on the whole real line but it is convex on the intervals $(0,\infty)$ and $(-\infty,\ 0)$. The key to everything here is that $\tilde{\theta}(x)$ has the same sign as x and this allows us to apply convexity of $\tilde{V}(x)-\frac{\rho}{2}x^2$ on each of the intervals $(-\infty,\ 0)$ and $(0,\infty)$ to conclude that

$$\begin{aligned}
\tilde{V}(x)-\tilde{V}(\tilde{\theta}(x))-\tilde{V}'(\tilde{\theta}(x))(x-\tilde{\theta}(x))&\geq\frac{\rho}{2}\left(x^2-\tilde{\theta}(x)\right)\\
-2\tilde{\theta}(x)(x-\tilde{\theta}(x))&=\frac{\rho}{2}(x-\tilde{\theta}(x))^2
\end{aligned}\tag{154}$$

From here we can follow word by word the proof of Theorem (6.3.7).

For the case $V(x)=rx$, we have equality in (152) if $\tilde{\theta}(x)=x+m\,sign(x)$ and thus this means $\theta(x)=\left(\sqrt{x}+m\right)^2$

In the case $V(x)=rx-s\log(x)$, we look at $\tilde{\theta}(x)=x+m$ for large m. In this case $\tilde{V}(x)=rx^2/2-s\log|s|$ and then a simple calculation shows that (152) is equivalent to

$$rm^2 + 2mr\int |x|\tilde{\mu}_v(dx) - 2s\int \log\left(\frac{|x+m\,sign(x)|}{|x|}\right)\tilde{\mu}_v(dx)$$
$$-2\int\int \log\left(1+m\frac{sign(x)-sign(y)}{x-y}\right)\tilde{\mu}(dx)\tilde{\mu}(dy)$$
$$\leq \int\left(r-\frac{s}{x(x+m\,sign(x))}\right)^2\tilde{\mu}_v(dx)$$

Dividing both sides by m^2 and taking the limit of m to infinity implies that $\rho \leq r$. On the other hand $\rho = r$ validates (152), hence $\rho = r$ is the best constant.

Next in line is the HWI inequality which is the content of the following statement.

Theorem (6.3.24) [177]: Assume V is as in Theorem (6.3.19) and the distance W given by (149). Then for any measure $\mu \in \mathrm{P}([0,\infty))$,

$$\mathrm{E}_V(\mu) - \mathrm{E}_V(\mu_V) \leq \sqrt{2I_V(\mu)}\mathrm{W}(\mu,\mu_V) - \rho\mathrm{W}^2(\mu,\mu_V) \tag{155}$$

For the case of $V(x) = rx - s\log(x)$, $r > 0$, $s \geq 0$, this inequality for $\rho = r$ is sharp.

Proof: As it was made clear in the previous two theorems, we translate this inequality in terms of the associated symmetric measures on $\mathbf{R}$. Following upon the proofs of above theorems, we can rewrite (155) in the following form:

$$\left(\int\left(H\tilde{\mu}(\theta(x)) - \tilde{V}'(\tilde{\theta}(x))\right)^2\tilde{\mu}_v(dx)\int\left(\tilde{\theta}(x)-x\right)^2\tilde{\mu}_v(dx)\right)^{1/2}$$
$$-\int\left(H\tilde{\mu}(\tilde{\theta}(x)) - \tilde{V}'(\tilde{\theta}(x))\right)\left(x-\tilde{\theta}(x)\right)\tilde{\mu}_v(dx)$$
$$+\int H\tilde{\mu}(\tilde{\theta}(x))\left(x-\tilde{\theta}(x)\right)\tilde{\mu}_v(dx) - \int\int \log\frac{x-y}{\tilde{\theta}(x)-\tilde{\theta}(y)}\tilde{\mu}_v(dx)\tilde{\mu}_v(dy) \geq 0$$

Using the fact that $\tilde{V}(x) - \frac{\rho}{2}x^2$ is convex on each interval $(-\infty, 0)$ and $(0,\infty)$ combined with the fact that x and $\tilde{\theta}(x)$ have the same sign, the rest of the proof is the same as the one of Theorem (6.3.8).

For the case $V(x) = rx - s\log(x)$, using $\theta(x) = \left(\sqrt{x}+m\right)^2$, one can show that $\rho = r$ is sharp.

At last, we would like to discuss a Poincaré type inequality in this context. As in this section, for the heuristics, we consider the general model of random matrices with distribution

$$\mathrm{P}_n(dM) = C_n e^{-nrTrM} (\det M)^{sn}\, dM = C_n e^{-nTr(rM - s\log(M))} dM = C_n e^{-nTr(V(M))} dM \quad (156)$$

where dM stands for the Lebesgue measure on $n \times n$ positive definite matrices and $s \geq 0$. For a given smooth compactly supported function $\phi : [0,\infty) \to \mathbf{R}$, we want to apply the Brascamp–Lieb inequality [95] to the function $\Phi(M) = T_r\phi(M)$ on the space of positive definite matrices.

Now, $\nabla\Phi(M) = \phi'(M)$.

The Hessian of $\Psi(M) := T_r(V(M))$ can be interpreted as a linear map from H_n(n ?n Hermitian matrices) into itself which is given by $\nabla^2\Psi(M)X = sM^{-1}XM^{-1}$. Hence the inverse of the Hessian is then $\left(\nabla^2\Psi(M)\right)^{-1} X = \frac{1}{s} MXM$. Thus we obtain from Brascamp-Lieb that

$$\int \frac{1}{n} T_r\left(\left(\nabla^2\Psi(M)\right)^{-1} \phi'(M)^2\right) \mathrm{P}_n(dM) \geq Var_{\mathrm{P}_n}(\Phi(M)).$$

On the other hand, from [36] or [258] the variance of $\Phi(M)$ converges to $\frac{1}{4}$ $Var_{\arcsin e_{[a,b]}}(\phi)$, where we recall that $arcsnie_{[a,b]} = \frac{dx}{\pi\sqrt{(x-a)(b-x)}}$ is the arcsine law on the support [a, b] of μ_V. Next,

$\frac{1}{n} T_r\left(\left(\nabla^2\Psi(M)\right)^{-1} \phi'(M)^2\right) = \frac{1}{sn} T_r\left(\left(\phi'(M)M\right)^2\right)$, whose integral against P_n converges to the integral of $\frac{1}{s} x^2\phi'(x)^2$ against the equilibrium measure μ_V from Eq. (143). These considerations suggest that

$$\int x^2\phi'(x)^2 \mu_V(dx) \geq \frac{s}{4} Var_{\arcsin e_{[a,b]}}(\phi) \quad (157)$$

Corollary(6.3.25)[274]: Show that

(i) $\mathrm{W}(\mu,\mu_V) = \frac{\sqrt{2I_V(\mu)}}{2\rho}$ (ii) $\mathrm{W}^2(\mu,\mu_V) = \frac{1}{2\rho^2}\int_0^\infty x\left(\int \frac{2}{x-y}\mu(dy) - V'(x)\right)^2 \mu(dx)$.

Proof:(i)Theorem (6,3.23) and(6.3.24) gives that

$$0 \le -\sqrt{2I_V(\mu)}\mathrm{W}(\mu,\mu_V) - \rho \mathrm{W}^2(\mu,\mu_V)$$

Hence

$$\mathrm{W}(\mu,\mu_V) = \frac{\sqrt{2I_V(\mu)}}{2\rho}$$

$$2\rho^2\mathrm{W}^2(\mu,\mu_V) = \frac{1}{2\rho^2}\int_0^\infty x\left(\int \frac{2}{x-y}\mu(dy) - V'(x)\right)^2 \mu(dx)$$

Hence the result.

Notice here that one can actually make this heuristic into an actual proof of this inequality.Motivated by these heuristics and also inspired by Theorem (6.3.14), we have the following stronger result.

Theorem (6.3.26) [177]: Assume that $Q:[0,\infty)\to\mathbf{R}$ is a convex potential and let $V(x)=Q(x)-s\log(x)$ for $s>0$ satisfy $Lim_{x\to\infty}(V(x)-2\log(x))=\infty$. Assume that the support of μ_V is[a, b]. Then for any smooth function ϕ on [a, b], the following holds

$$\int x^2\phi'(x)^2\mu_V(dx) \ge \frac{s}{4\pi^2}\int_a^b\int_a^b\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2\left(\frac{-2ab+(a+b)(x+y)-2xy}{2\sqrt{(x-a)(b-x)}\sqrt{(y-a)(b-y)}}\right)dxdy \quad (158)$$

If $Q(x)=rx+t$, equality is attained for $\phi(x)=C_1+\frac{C_2}{x}$ φ(x), therefore (158) is sharp.

In particular, combining (158) with (140) for $\lambda=0$, we get an improvement of (147) as

$$\int x^2\phi'(x)\mu_V(dx) \ge \frac{s}{2}Var_{aecsine_{[a,b]}}(\phi)$$

Equality though is attained only for φ identically 0.

In the case $V(x)=rx,\ \ r>0$, on $[0,\infty)$,there is no constant $C>0$ such that inequality (158) holds with C instead of $\frac{s}{4\pi^2}$. Nevertheless, for every smooth ϕ on[a, b], the following holds,

$$\int x\phi'(x)^2\mu_V(dx) \ge \frac{r}{4\pi^2}\int_a^b\int_a^b\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2\left(\frac{-2ab+(a+b)(x+y)-2xy}{2\sqrt{(x-a)(b-x)}\sqrt{(y-a)(b-y)}}dxdy\right) \quad (159)$$

with equality for $\phi(x)=C_1+C_2x$.

As remarked after the statement of Theorem (6.3.14) the numerator in (159) is nonnegative.

Proof: The same argument as in the proof of Theorem (6.3.14), shows that the density $g(x)$ of μ_V satisfies

$$g(x)\geq\frac{s\sqrt{(x-a)(b-x)}}{2\pi x\sqrt{ab}},$$

therefore it suffices to show that

$$\frac{1}{\pi\sqrt{ab}}\int_a^b x\phi'(x)^2\sqrt{(x-a)(b-x)}dx\geq\frac{1}{2\pi^2}\int_a^b\int_a^b\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2 \times\frac{-2ab+(a+b)(x+y)-2xy}{2\sqrt{(x-a)(b-x)}\sqrt{(y-a)(b-y)}}dxd$$

Next, making the change of variable $x=(a+b)/2+u(b-a)/2$ and denoting

$$\zeta(u)=\phi((a+b)/2+u(b-a)/2),$$

We reduce the problem to showing that for any smooth function ϕ on$[-1,\ 1]$, we have

$$\frac{1}{\pi\sqrt{an}}\int_{-1}^{1}\left(\frac{a+b}{2}+\frac{a-b}{2}u\right)\zeta(u)^2\sqrt{1-u^2}du\geq\frac{1}{2\pi^2}\int_{-1}^{1}\int_{-1}^{1}\left(\frac{\zeta(u)-\zeta(v)}{u-v}\right)^2 \times\frac{1-uv}{\sqrt{1-u^2}\sqrt{1-v^2}}dudv$$

Denoting $\beta=\frac{b-a}{b+a}$, we have that $\frac{a+b}{2\sqrt{ab}}=\frac{1}{\sqrt{1-\beta^2}}$, and the preceding inequality reformulates as

$$\int(1+\beta u)\zeta'(u)^2\sqrt{1-u^2}du\geq\frac{\sqrt{1-\beta^2}}{2\pi}\int_{-1}^{1}\int_{-1}^{1}\left(\frac{\zeta(u)-\zeta(v)}{u-v}\right)^2\left(\frac{1-uv}{\sqrt{1-u^2}\sqrt{1-v^2}}\right)dudv \qquad (160)$$

To show this, take $\psi(t)=\zeta(\cos(t))$ and then after the change of variable $u=\cos(t)$ we need to check

$$\int_0^{\pi}\left(1+\beta\cos(t)\right)\psi'(t)^2\,dt \ge \frac{\sqrt{1-\beta^2}}{2\pi}\int_0^{\pi}\int_0^{\pi}\left(\frac{\psi(t)-\psi(s)}{\cos(t)-\cos(s)}\right)^2\left(1-\cos(t)\cos(s)\right)dtds.$$

Writing $\psi(t)=\sum_{n=0}^{\infty}a_n\cos(nt)$ and using that $\psi'(t)=\sum_{n=0}^{\infty}na_n\sin(nt)$, together with the fact that

$$\int_0^{\pi}\cos(t)\sin(nt)\sin(mt)=\begin{cases}\frac{\pi}{4} & for\ |m-n|=1,\\ 0 & otherwise\end{cases}$$

and Eq. (153), the inequality becomes

$$\sum_{n\ge1}\left(n^2a_n^2+\beta n(n+1)a_na_{n+1}\right)\ge\sqrt{1-\beta^2}\sum_{n\ge1}na_n^2 \tag{161}$$

Let $\delta=\frac{1-\sqrt{1-\beta^2}}{\beta}$ be the solution $0<\delta<1$ of $\beta\delta^2-2\delta+\beta=0$. Notice that for any $n\ge1$, we have $a_na_{n+1}\ge-\frac{\delta}{2}a_n^2-\frac{1}{2\delta}a_{n+1}^2$

which implies that

$$\sum_{n\ge1}\left(n^2a_n^2+\beta n(n+1)a_na_{n+1}\right)\ge\sum_{n\ge1}\left(na_n^2-\frac{\beta n(n+1)}{2}\left(\delta a_n^2+\frac{1}{\delta}a_{n+1}^2\right)\right)$$
$$=\sum_{n\ge1}\frac{n\beta\left(1-\delta^2\right)}{2\delta}a_n^2=\sqrt{1-\beta^2}\sum_{n\ge1}na_n^2$$

what we had to prove. Notice here that equality is attained in this inequality if and only if $a_{n+1}=-\delta a_n$ for all $n\ge1$, which means that $a_n=(-1)^{n-1}\delta^{n-1}a_1$. This corresponds to the function $\psi(t)=a_1\frac{\delta+\cos(t)}{1+\delta^2+2\delta\cos t}$, or $\zeta(u)=a_1\frac{\delta+u}{1+\delta^2+2\delta u}$

which means that $\phi(x)=a_1(r-s/x)$. Therefore equality holds also for $\phi(x)=C_1+C_2/x$.

For the second part, in the case $V(x)=rx$ with $r>0$, notice that if there is a $C>0$ so that (158) holds with C instead of $s/4\pi^2$, then, following the same argument as above, we would have the equivalent of (161) as

$$\sum_{n\ge1}\left(n^2a_n^2+n(n+1)a_na_{n+1}\right)\ge C\sum_{n\ge1}na_n^2$$

Taking in this $a_n = \frac{(-r)^n}{n}$ for $0 < \gamma < 1$, we have that $\gamma^2 / (\gamma+1) \geq -C\log(1-\gamma^2)$, and this is certainly false for γ close to 1.For Eq. (159), notice that in this case the equilibrium measure is $\mu_V(dx) = \frac{r\sqrt{b-x}}{2\pi\sqrt{x}}$ and then after a simple rescaling this follows from Eq. (120). This complete the proof of the theorem.

It is interesting to look at this inequality as a spectral gap result. For example in the case of the Marcenko–Pastur measure $(Q(x) = rx)$, the inequality (158) is actually equivalent to inequality (160). Using the interpretation from end of this section, we can rephrase this as, for a given $\beta \in (0,1)$,

$$\int (1+\beta x)(1-x^2)\phi'(x)^2 v_0(dx) \geq \sqrt{1-\beta^2}\,(N\varphi, \phi)$$

where v_0 is the arcsine law on $[-1, 1]$ and N is the number operator. Now we can define the operator

$$L_\beta\phi(x) = -(1+\beta x)(1-x^2)\phi''(x) - (\beta - x - 2\beta x^2)\phi'(x)$$

With this definition,

$$\langle L_\beta\phi, \phi\rangle_{v_0} = \frac{1}{\pi}\int (1+\beta x)\phi'(x)^2\sqrt{1-x^2}\,dx$$

and then inequality (160) becomes

$$\langle L_\beta\phi, \phi\rangle \geq \sqrt{1+\beta^2}\,\langle N\phi, \phi\rangle_{v_0}$$

for any smooth function ϕ on $[-1, 1]$. In particular this means that $L_\beta \geq \sqrt{1-\beta^2}N$. On the other hand it is clear that the operator L_β can not be diagonalized by the Chebyshev polynomials of the first kind, therefore the orthogonal polynomial approach given does not work the same way here.

List of symbols

symbol		Page
vol	Volume	1
$\inf$	Infimum	1
Lip	Lipschitz	3
Ric	Ricci	3
$\exp$	Exponential	3
det	Determinant	5
L^{∞}	Lebesque measure	7
$Hess$	Hessian	7
$dist$	Distant	7
$sign$	Signature	7
$\sup$	Supremum	15
dim	Dimension	17
W_p	Free Wasserstein metric	19
$\otimes$	Tensor product	22
$\boxplus$	Operation	24
Im	Imajnary	24
supp	Support	25
L^3	Banach space	26
L^2	Hilbert space	30
min	Minimum	44
$diam$	Diameter	52
$a\ c$	Absolutely continuous	57
LSI	Logarthmic Solev Inequality	85
Tr	Trace	90
$curv$	Cuevature	108
max	Mascimum	112
$\lvert . \rvert$	Cardinality	124
ℓ^{∞}	Lebesgue space	175
T^{N}	Abelian group	180
Var	Variation	227
SG	Spectral Gap	228
SP	Super poincaire inequality	230
LO	Lata la-Oleszkiewicz inequality	236
a.e	Almost every where	241
sgn	Signature	244
$\oplus$	Direct sum	247
IC	Inf.convolution	253
TC	Transportation cost	254
Cap	Capacity	260
opt	Optimal	260
$\mathrm{L}(\mu, \mathrm{V})$	Levydistance	293

References

[1] A. A. Borovkov, S. A. Utev, An inequality and a characterization of the normal distribution connected with it. Teor. Veroyatnost. i Primenen., 28(2), 209-218 (1983

[2] A. Arnold, P. Markowich, G. Toscani and A. Unterreiter, on logarithmic Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Plank Type equation . Comm. Partial Differential Equations, to appear .

[3] A. Connes, Compact metric spaces, Fredholm modules and hyperfinite-ness, Ergodic Theory and Dynamical Systems 9 (1989),207-220

[4] A. D.A. Lexandrov, A theorem on triangles in a metric space and some applications.Trudy mat.Inst.Steklov, 38(1951)5-23 (Russian translated into German and combined with more material in [2]),

[5] A. Dembo and O. Zeitouni. Large deviatians techniques and applicatians. Secand editian. Applications of Mathematics, 38. Springer Verlag, (1998).

[6] A.D. Alexandrov, Über eine Verallgemeinerung der Riemannschen Geometrie. Schr. Forschungsinst.Math. Berlin, 1 (1957), 33-84.

[7] A.I. Bonciocat, K.T. Sturm, Mass Transportation and rough curvature bounds for discrete spaces. APP. Math. Geom. Univ. of Bonn 82/1, 53115,(2009)

[8] A. Kasue, & H. Kumura, Spectral convergence of Riemannian manifolds. Tohoku Math, J., 46 (1994), 147-179

[9] A. Kasue, Convergence of Riemannian manifolds and Laplace operators II. Preprint, 2004.

[10] A. Khintchine, Über dyadische Brűche, Math. Z.18 (1923),109-116. (German).

[11] A.M. Khorunzhy, B.A. Khoruzhenko, L. Pastur, Asymptotic properties of large random matrices with independent entries, J. Math. Phys. 37 (10) (1996) 5033–5060.

[12] A.Nica, D. Shlyakhtenko, R. Speicher, Some minimization problems

for the free analogue of the Fisher information. Adv. Math., 141(2), 282-321 (1999).

[13]A. Ronami, Ftude des coefficients de Fourier des functions de L^p (G), Ann. Inst, Fourier (Grenoble) 20 (1970) 335-402.

[14] A. W. Knapp, Representatior, Theory of Semisimple Groups, An Overview Based on Examples, Princeton University Press, Princeton, 1986.

[15] A. W. Marshall and I. Olkin, Inequalities: Theory of majorization and its applications, Academic Press, New York, 1979.

[16] B. Efron, Student's t-test under symmetry conditions, J. Amer. Statist. Assoc. 64 (1969), 1278-1302.

[17] B. Maurey. Some deviation inequalities. Geom. Funct. Anal., 1(2):188-197, (1991).

[18] B. Muckenhoupt. Hardy's inequality with weights. Studia Math., 44:31–38, 1972.

[19] B. Simon and R , Hoegh-Krohn , "Hypercontractive semigroups and two dimensional self coupled Bose fields" J. Functional Anal, 9 (1972), 121-180

[20] C. Ane, S. Blachere, D. Chafaı, P. Fougeres, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer. Sur les inegalit´es de Sobolev logarithmiques, volume 10 of Panoramas et Synth`eses [Panoramas and Syntheses. Socieete Math´ematique de France, Paris, (2000).

[21] C. Plaut, Metric spaces of curvature ≥ k, in Handbook of Geometric Topology, pp. 819-898. North-Holland, Amsterdam, (2002).

[22] C. Villani, Optimal transport: old and new, Grundlehren Der Mathematischen Wissenschaften 338 (2008), Springer.

[23] C. Villani. Topics in Optimal Transportation. Graduate Studies in Mathematics 58. American .Matheruatical Society, Providence RI, (2003).

[24] C.C. Graham , and O.C. McGehee , Easy in commutative Harmonic analysis Springer-Verlag, New York, (1979).

[25] D. Bakry, "L'hypercontractivive et son utilisation en .theorie des

semigroupes”, in -Lectures on probabIlIty theory (Samt-Flour, 1992), Lecture Notes in Math. 1581, Spnnger, Berlm, p.1-114 (1994).

[26] D. Bakry, F. Barthe, P. Cattiaux, and A. Guillin. A simple proof of the Poincar´e inequality for a large class of probability measures including the log-concave case. Electron. Commun. Probab., 13:60–66, 2008.

[27] D. Bakry, M. Emery, Diffusions hypercontractives. Seminaire de Probabilitts, XIX, 1983/84, 177-206, Lecture Notes in Math., 1123, Springer, Berlin, 1985.

[28] D. Bakry, M. Emery, Hypeconcentrative de semi-groups de diffusion. C. R. Acad. Sci. Paris.Ser. 1 math. 299(1987).no.15,775-778.

[29] D. Burago, Y. Burago, S. Ivanov, A course in metric geometry. Graduate Studies in Mathematics, 33. American Mathematical Society, Providence, R.I., 2001.

[30]D. Cordero. Erausquin, Sur Ie transport de measures periodiques, C.R. A cad. Sci. Paris I 329 (1999), 199-202 .

[31]D. Cordero-Erausquin, R.J. McCann and M. Schmuckenschläger, “A Riemannian interpolation inequality it la Borell, Brascamp and Lieb”, Invent. Math. 146, p. 219-257 (2001).

[32]D. Cordero-Erausquin, Some applications of mass transport to Gaussian-type inequalities, Arch. Ration. Mech. Anal. 161 (3) (2002) 257–269.

[33]D. Cordero-Erausquin, W. Gangbo, and C. Houdre. Inequalities for generalized entropy and optimal transportation. In Recent advances in the theary and applicatians af mass transpart, volume.253 of Cantemp. Math., pages 73-94. Amer. Math. Soc., Providence, RI, (2004).

[34] D. Feyel, & A. S. ÜstÜhel, Monge-Kantorovitch measure transportation Monge-Ampere equation on Wiener space. Probab. Theory Related Fields, 128 (2004), 347-385.

[35] D.Feyel, and A.S.ÜstÜnel ,The strong solution of the Monge-Ampere equation on the Wiener space for log-concave densities. C. R. Math. Acad. Sci. Pans, 339 (2004), 49-53.

[36]D. Jonsson, Some limit theorems for the eigenvalues of a sample covariance matrix, J. Multivariate Anal. 12 (1) (1982) 1–38.

[37]D. Rider, Randomly continuous functions and Sidon sets, Duke Math. J., 42 (1975), 759-764.

[38]D. Shlyakhtenko, Free fisher information with respect to a complete positive map and cost of equivalence relations, Comm. Math. Phys. 218 (2001), 133-152

[39]D. V. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory..11. Invent. Math., 118(3), 411-440 (1994).

[40]D. V. Voiculescu, Cyclomorphy. MSRI preprint # 2001-019.

[41] D. V. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory IV. Maximum entropy and freeness. Free probability theory (Waterloo, ON, 1995), 293-302, Fields Inst. Commun. 12, Amer., Math. Soc., Providence, RI (1997).

[42]D. Voiculescu, K. Dykema, A. NICA, Free Random Variables, CRM Monograph Series no. 1, American Mathematical Society, Providence, RI,1992.

[43]D. Voiculescu, The analogue of entropy and of Fisher's information measure in free probability theory VI: Liberation and mutual free information, Adv. Math. 146 (1999), 101-166.

[44]D. Voiculescu, The analogues of entropy and of Fisher’s information measure in free probability theory. V. Noncommutative Hilbert transforms, Invent. Math. 132 (1) (1998) 189–227.

[45]D. Voiculescu, The analogues of entropy and of Fisher’s information measure in free probability theory. I, Comm.Math. Phys. 155 (1) (1993) 71–92.

[46]D.Bakfy, & Z. Qian, Some new results on eigenvectors via dimension, diameter, and , Ricci curvature. Adv. Math., 155 (2000), 98-153

[47] D. Voiculescu, The derivate of order ½ of a free convolution by a semicircle distribution, Indiana Univ. Math. J 46 (1997), 697-703 .

[48] E,- Y. Wang, On estimation of the logarithmic Sobolev constant and gradient estimates of heat semigroups. Probab. Theory Related Fields 108 (1997). no. 1,87-101.

[49]E. B. Davies, Heat Kernels and Spectral Theory. Cambridge Tracts in Mathematics, 92 Cambridge University Press, Cambridge, 1989.

[50] E. B. Davies, J. M. Lindsay, Noncommutative symmetric Markov semigroups. Math. Z., 210(3), 379-411 (1992).

[51] E. B. Saff and V. Totik, Logarithmic Potentials with External Fields, Springer-Verlag, Berlin-Heidelberg New York, (1997).

[52]E. B. Saff and V. Totik, Logarithmic Potentials with External Fields, Appendix B by T. Bloom, Grundlehren der Mathematischen Wissenschaften, 316, Springer-Verlag, Berlin (1997).

[53] E. M. Stein , “Interpolation of linear operators”, Trans Amer. Math. Soc., 83 (1956), 482-492.

[54] E. Nelson, Topica In Dynamics I: Flow, Princeton University Press, (1970).

[55] E. Nelson, "A quartic interaction in two dimensions," Mathemtical Theory of Elementary Particles, R. Goodman and I. E. Segal, Ed., M.T. Press, 1966, Cambridge, Mass.

[56] E. Nelson, “ The free Markoff field”, J; of Functional Anal., 12 (1973), 211-227.

[57]F. Barthe and C. Roberto. Modified logarithmic Sobolev inequalities on R. Potential Anal., 29(2):167–193,(2008).

[58] F. Barthe and C. Roberto. Sobolev inequalities for probability measures on the real line. Studio.Math., 159(3):481-497, (2003).

[59]F. Barthe, P. Cattiaux, and C. Roberto. Interpolated inequalities between exponential and Gaussian, Orlicz hypercontractivity and isoperimetry. Rev. Mat. Iberoamericana, 22(3):993–1067, (2006).

[60] F. Barthe, P. Cattiaux, and C. Roberto. Isoperimetry between exponential and Gaussian. Electron. J.Probab., 12:no. 44, 1212–1237 (electronic), (2007).

[61] F.Barthe and C. Roberto. Modified Logarithmic-Sobolev inequalities. Preprint, (2006).

[62]F. Hiai and D. Petz, A large deviation theorem for the empirical eigenvalue distribution of random unitary matrices, Ann. Inst. H. Poincare Probab. Statist. 36 (2000), 71-85.

[63]F. Hiai and D. Petz, Properties of free entropy related to polar decomposition, Comm. Math. Phys. 202 (1999), 421-444.

[64]F. Hiai, D. Petz, Eigenvalue density of the Wishart matrix and large deviations, Infin. Dimens. Anal. Quantum Probab. 1 (1998) 633–646. [287]

[65]F. Hiai, D. Petz, The Semicircle Law, Free Random Variables and Entropy, Math. Surveys and Monographs, vol. 77, American Mathematical Society, (2000).[154]

[66]F. Hiai, D. Petz, Y. Ueda, Free transportation cost inequalities via random matrix approximation, Probab. Theory Related Fields 130 (2004)199-221.

[67] F. Hiai, D. Petz, Y. Veda, Inequalities related to nee entropy derived nom random matrix approximation (2003).

[68] F. Hiai, M. Mizuo and D. Petz, Free relative entropy for measures and a corresponding perturbation theory, J. Math. Soc. Japan 54 (2002), 679-718.

[69] F. Otto and C. Villani, "Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev inequality", J. Funct. Anal. 173, p. 361-400 (2000).

[70] F. Otto, T le geometry of dissipative evolution equations: the porous medium equation, Comm. Partial Differential Equations, in press.

[71]F.Y. Wang, Functional Inequalities, Markov Properties and Spectral Theory, Science Press, Beijing–New York,(2005).

[72]F-Y Wang. Probability distance inequalities on Riemannian manifolds and path spaces. .J. Funct. Anal., 206(1):167-190,(2004).

[73]F-Y. Wang. A generalization of Poincare and log-Sobolev inequalities. Potential Analysis, 22:1–15, (2005).

[74]F-Y.Wang. Generalized transportation-cost inequalities and applications. Potential Anal., 28(4):321–334,2008.

[75]F-Y. Wang. Functional inequalities for empty essential spectrum. J. Funct. Anal., 170(1):219–245, (2000).

[76]G. Da Prato, M. Rockner, Singular dissipative stochastic equations in Hilbert spaces. Probab.'Theory Related Fields 124 (2002), no. 2, 261-303

[77] G. Alberti and L. Ambrosio, " geometrical approach to monotone functions in $\mathbf{R}^{nn}$, Math. Z. 230, p.259-316 (1999).

[78]G. Ben Arous, A. Guionnet, Large deviations for Wigner's law and Voiculescu's non-commutative entropy, Probab.Theory Related Fields 108 (2) (1997) 183–215.

[79]G. Ben Arous, A. Guionnet, Large deviations for Wigner's law and Voiculescu's noncommutative entropy, Probab. Theory Related Fields 108 (1997) 517-542.

[80]G. Blower, The Gaussian isoperimetric inequality and transportation, Positivity 7 (3) (2003) 203–224.

[81]G. Buttazzo, M. Giaquinta and S. Hildebrandt, One-dimensional variational problems, Oxford Lecture Series in Mathematics and its Applications 15, Oxford University Press, New York (1998).

[82]G. Feissner, "A Gaussian measure analogue to Sobolev's inequality," Ph.D. Thesis, Cornell Univ. (1972).

[83] G. Feissner, "Hypercontractive semigroups and Sobolev's inequality," Tran. Amer. Math. Soc. 210 (1975), 51-62.

[84]G. Johnson, G. Woodward, On p-Sidon sets, Indiana Univ. Math. J. 24 (1974) 161–167.

[85]G. Monge, Mémoire Sur la théorie desn déblais et desreblais, in Histoire de l'Académie Royale des Sciences Paris, 1781.

[86]G. Perelman, The entropy formula for the Ricci flow and its geometric applications.Preprint, (2002).

[87]G. Peskir, Best constants in Kahane-Khintchine inequalities for complex Steinhaus function, Proc. Amer. Math.Soc. 123 (10) (October 1995) 3101–3111.

[88] G. Petkir, Best constants in Kahane-Khintchine inequalities in Orlicz spaces, Institute of Mathematics, University of Aarhus, Preprint Series No. 10, 1992 (42 pp.); J. Multivariate .Anal. 45 (1993), 183-216.

[89] G. Pisier, Sur l'espace de Banach des series de Fourier aleatoires presque surement continues, Se'minaire sur la Geometrie des Espaces de Banach 1977-1978 Exposes No. XVII-XVIII.

[90]G. W. Johnson and G. S. Woodward, On p-Sidon sets, Indiana Univ. Math. J., 24 (1974), 161-167.

[91]G.N.Watson, Notes on generating functions of polynomials ,(3) polynomials of Legendre and Gegenbauer, J. London Math. Soc. 8 (1933) 289–292.

[92]H, Bercovici, D. Voiculescu, Free convolution of measures with un-bounded support, Indiana Univ. Math. J. 42:3 (1993), 733-773.

[93]H. Djellout, A. Cuillin, and L. Wu. Transportation cost-information inequalities and applications to random dynamical systems and diffusions. Ann. Probab., 32(3B):2702-2732, 2004.

[94]H. F. Trotter, “Approximation of semigroup of operator,” Pacific J. Math., 8 (1958), 887-919.

[95]H.J. Brascamp, E.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, J. Funct. Anal. 22 (4) (1976) 366–389.

[96]I. Csiszár, Information-type measures of difference of probability distributions and indirect observations, Studia Sci. Math. Hungar. 2 (1967) 299–318.

[97]I. Chavel, Riemannian Geometry-a Modern Introduction. Cambridge Tracts in Mathematics, 108. Cambridge University Press, Cambridge, 1993.

[98]I. Dumitriu and A. Edelman, Matrix models for Beta Fnsembles,

Journal of Mathematical physics 43 (2002), no. 11, 5830-5847.

[99]I. E. Segal, “Constriction of nonlinear local quantum processes II ,”Inventiones Math., 14 (1971),211-241.

[100] I. E. Segal, “Construction of nonlinear local quantum processes I Annals of Math., 92 (1970).462-481.

[101]I. Gentil, A. Guillin, and L. Miclo. Modified logarithmic Sobolev inequalities and transportation inequalities.Probab. Theory Related Fields, 133(3):409–436, 2005.

[102]I. Popescu, Talagrand inequality for the semicircular law and energy of the eigenvalues of beta ensembles, Math.Res. Lett. 14 (6) (2007) 1023–1032.

[103]I. Gentil. From the Prekopa-Leindler inequality to modified logarithmic Sobolev inequality. Annales de la Facult´e des Sciences de Toulouse, 17(2):291–308, 2008.

[104]J. Heinonen, Lectures on Analysis on Metric Spaces. Springer, New York, 200l.

[105] J. Kigami, Analysis on Fractals. Cambridge Tracts in Mathematics, 143. Cambridge Uni~ l.versjty Press, Cambridge, 2001.

[106] J. A. Cuesta-Albertos and A. Tuero-Diaz. A characterization for the solution of the Monge Kantorovich mass transference problem. Statist. Probab. Lett., 16(2):147-152, (1993).

[107]J. Cheeger and T. Colding, “On the structure of spaces with Ricci curvature bounded below III”, J. Diff. Geom. 54, p, 37-74 (2000) .

[108]J. Cheeger and T. Colding, “On the structure of spaces with Ricci curvature bounded below I”, J. l-Diff. Geom. 46, p. 37-74 (1997).

[109]J.Cheeger, & COLDING, T. H., On the structure of spaces with Ricci curvature bounded below I, II, III. J- Differential Geom., 46 (1997), 406—480; Ibid, 54 (2000), 13--35; Ibid,54 (2000), 37-74.

[110] J.Cheeger, Differentiability of Lipschitz functions on metric measure spaces. Geom. Funct. ,Anal. 9 (1999), no. 3,428-517

[111] J. D. DeuscheI and D. W. Stroock, Large deviations, AMS revised edition (1989).

[112] J. E. Littlewood, On bounded bilinear forms in an infinite number of variables, Quart.J. Math. Oxford, 1 (1930), 164-174.

[113] J. Glimm, “Boson fields with nonlinear self interaction in two dimensions,” Commun. Math.Phys., 8 (1968), 12-25.

[114] J. Hoffmann-Jorgensen, Function norms, Institute of Mathematics, University of Aarhus, Preprint Series No. 40, 1991.

[115]J. L. Sauvageot, Quantum Dirichlet forms, differential calculus and semigroups. Quantum probability and applications, V (Heidelberg, 1988),334-346, Lecture Notes in Math., 1442, Springer, Berlin (1990).

[116] J. Lindenstrauss, L. Tzafriri, Classical Banach Spaces I, Springer-Verlag, Berlin, Heidelberg, New York, (1977).

[117] J. Lott and C. Villani. Hamilton-Jacobi semigroup on length spaces and applicotions. J. Math.Pures Appl. (9), 88(3):219-229, (2007).

[118]J. Lott, &C. Villani, Weak curvature conditions and Poincare inequalities. Preprint, 2005.

[119] J. Lott, “Some geometric properties of the Bakry-Emery-Ricc tensor”, Comment. Math. Helv. 78, p. 865-883 (2003).

[120] J. Lott, C. Villani, Ricci curvature for metric-measure spaces via optimal transport, to appear in Ann. of Math. (2005).

[121] J. Lott, C. Villani, Ricci curvature for metric-measure spaces via optimal transport, to appear in Ann. of Math. (2009).

[122] J. Milnor, Curvature of left invariant metrics on Lie groups, Adv. Math., 21 (1976), 293-329.

[123] J. Urbas, On the second boundary value problem for equations of Monge-Ampere type, J. Reine Angew. Math. 487 (1997), 115-124.

[124] J.D. Benamou and Y. Brenier, A numerical method for the optimal time-continuous mass transport problem and related problems, in "Monge. Ampere Equation: Applications to Geometry and Optimization," Con temp.

Math., Vol. 226, pp. 1-11, Arner. Math. Soc., Providence, RI, (1999).

[125] J.D. Deuschel and D.W. Stroock. Large deviatians, volume 137 of Pure and Applied Mathematics. Academic Press Inc., Boston, MA, 1989.

[126] J.H.B. Kemperman, On the optimum rate of transmitting information, Ann. Math. Statist. 40 (1969) 2156–2177.

[127] J.-P. Kahane, Some Random Series of Functions, Heath Math. Monographs, Mass.,(1968).

[128] K. Fukaya, Collapsing of Riemannian manifolds and eigenvalues of Laplace operator, Invent. IVlath. 87 (1987) 517--547.

[129] K. Grove and P. Petersen, "Manifolds near the boundary of existence", J. Diff. Geom. 33, p. 379-394 (1991) .

[130] K. Ito, "Complex multiple Wiener integral." Japanese J. of Math., 22 (1953), 63-86.

[131]K. Ito, "Multiple Wiener integral," 1. of Math. Soc. of Japan, 3 (1951), 157-169.

[132] K. Johansson, On fluctuations of eigenvalues of random Hennitian matrices, Duke Math. J. 91 (1998) 151-204.

[133] K. Kuwie, & T. Shioya, On generalized measure contraction property and energy functicnals over Lipschitz maps. Potential Anal., 15 (2001), 105-12l.

[134] K.Kuwae, & T. Sgioya, Convergence of spectral structures: a functional analytic theory and its applications to spectral geometry. Comm. Anal. Geom., 11 (2003),599-673.

[135] K. Kuwae and T.Shioya and P.Koskela,Sobolev and Dirichlet spaces over maps between metric spaces. J. Reine Angew. Math.,555 (2003), 39-75.

[136] K. Marton, A measure concentration inequality for contracting Markov chains, Geom. Funct. Anal. 6 (1997) 556-571.

[137] K. Marton, A measure concentration inequality for contracting Markov chains, Geom. Funct. Allal. 6 (1996), 556-571.

[138] K. Marton. A simple proof of the blowing-up lemma. IEEE Trans. Inform. Theory, 32(3):445-446,(1986).

[139] K. Marton. Bounding d-distance by informational divergence: a method to prove measure concentration.Ann. Probab., 24(2):857–866, 1996.

[140] K. T. Sturm, A semigroup approach to harmonic maps. Potential Ana/., to appear. Available online at: http://sp611-iam.uni.bonn.de/publikation php?=39.

[141] K. T. Sturm, on the geometry of metric measure spaces. I, Acta Math. 196 (2006) 65-131.

[142] K.T. Sturm, Diffusion processes and heat kernels on metric spaces. Ann. Probab., 26 (1998), no.1,1-55.

[143] K.T. Sturm, Probability measure:, on metric spaces of nonpositive curvature, in Heat Kernels and Analysis on Manifolds, Grouphs, and Metric Spaces (Paris, 2002), pp. 357-390. Contemp. Math., 338. Amer. Math. Soc., Providence, RI, (2003).

[144] K.T. Sturm,Analysis on local Dirichlet spaces. III. The parabolic Harnack inequality..J. Math. Pures Appl., 75 (1996),273-297.

[145] K.t. Sturm, Convex functionals of probability measures and nonlinear diffusions on manifolds. J. Math. Pures Appl., 84 (2005), 149-168.

[146]K.T. Sturm, Generalized Ricci bounds and convergence of metric measure spaces. C. R. Math. Acad. Sci. Paris, 340 (2005), 235-238.

[147] K.T. Sturm, Metric spaces of lower bounded curvature. Exposition. Math., 17 (1999), 35-47.

[148] K.T. Sturm, A curvature-dimension condition for metric measure spaces. C. R. Math. Acad. Sci. Paris, 342 (2006), 197-200.

[149] K-T. Sturm, "On the geometry of metric measure spaces", to appear, Acta Math. 196(2006).133-177.

[150] L. Nirenberg, "On elliptic partial differential equation,"Ann, Scuola Normale Superiore di Pisa, Ser III, 13 (1959), 1-47.

[151] L. A. Caffardli, Boundary reularity of maps with convex potentials,

Comm. Pure Appl. Math.45(1992).1141-1151.

[152] L. A. Caffarelli, The regularity of mappings with a convex potential, J. Amer. Math. Soc.5 (1992), 99.104.[

[153]L. C. Evan; and R. F. Gariepy, Measure theory and fine properties of functions, in Studies in advanced Mathematics, CRC Press, Boca Raton, FL, 1992.

[154] L. Gross, "Existence and uniqueness of physical ground states," 1. Functional Anal., 10 (1972), 52-109.

[155] L. Gross, Logarithmic Sobolev inequalities, Amer. J. Math. 97 (4) (1975) 1061–1083.

[156] L. Mbrosio, & P. Tilli, Topics on Analysis in Metric Spaces. Oxford Lecture Series in Mathematics and its Applications, 25. Oxford University Press, Oxford, 2004.

[157] L. N. Wasserstein [Vasershtein], Markov processes over denumerable products of spaces describing large system of automata. Problemy Peredaci Informacii, 5:3 (1969),64-72 (Russian). English translation in Problems of Information Transmission, 5:3 (1969), 47-52.

[158] L. Saloff-Cosie, Convergence to equilibrium and logarithmic Sobolev constant on Manifold with Ricci curvature bounded below, Colloq.Math.67 (1994), 109-121.

[159] L. Saloff-Coste, Aspects of Sobolev-type Inequalities. London ,Math. Soc. Lecture Note Series, 289. Cambridge University Press, Cambridge, (2002).

[160] L. V. Kantoroich, & G. S. Rubinshtein, On a functional space and certain extremum problems. Dokl. Akad. Nauk SSSR, 115 (1957), 1058-1061.

[161] L. V. Kantorovich. On the translocation of masses. C. R. (Doklady) Acad. Sci. URSS, 37 (1942), 199-201.

[162] M. Delpino, & J. Dolbeault, Best constants for Gagliardo-Nirenberg inequalities and applications to nonlinear diffusions. J. Math. Pures Appl., 81 (2002), 847-875.

[163] M. Gromov, Groups of polynomial growth and expanding maps. Inst. Hautes Etudes Sci. Publ. Math. (1981), 53-73.

[164] M. Gromov, Hyperbolic groups. Essays in group theory, 75-263, Math. Sci. Res. Inst. Publ. 8 (1987), Springer, New York.

[165] M. Gromov and V.D. Milman. A topological application of the isoperimetric inequality. Amer. J. Math.,105:843–854, (1983).

[166] M. Gromov, Metric structures for Riemannian and non-Riemannian spaces, Progress in Mathematics 152, Birkhäuser, Boston (1999).

[167] M. Kac, “Random walk and the theory of Brownian motion,” Amer. Math. Monthly, 55 .(1947), 369-391.

[168] M. L. Mehta, Random Matrices, Second edition, Academic Press, Boston, (1991).

[169] M. Ledoux On an integral criterion for hypercontractivity of diffusion semigroups and extremal function, J. Funct. Anal. 105 (1992), 444-465.

[170] M. Ledoux The geometry of Markov processes, lectures, Zurich, 1998.[112]

[171] M. Ledoux, A (one-dimensional) free Brunn–Minkowski inequality, C. R. Acad. Sci. Paris 340 (2005) 301–304.

[172] M. Ledoux, Concentration of measures and logarithmic Sobolev inequalities, In Seminaire de ProbabiliMs XXXIII, Lecture Notes in Math., Vol. 1709, Springer-Verlag, 1999, pp. 120---216.

[173] M. Ledoux, Measure concentration, transportation cost, and functional inequalities, Summer School on Singular Phenomena and Scaling in Mathematical Models, Bonn, (2003).

[174]M. Ledoux, The Concentration of Measure Phenomenon, Math. Surveys Monogr., vol. 89, Amer. Math. Soc., Providence,RI, (2001).

[175] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities, Lectures, Berlin, (1997)

[176] M. Ledoux. On Talagrands’s devation inequalities for product measure ESAIM probabstatist ., 1:1:63-87,(1996).

[177] M. Ledoux, I. Popescu, Mass Transportation proofs of free functional inqualities, and free poincare Inequalities, J.Funet . Anal. 257 (2009) 1175-1221

[178] M. R. Bridson, & A. Haefliger, Metric Spaces of Non-positive Curvature. Grundlehren /der Mathematischen Wissenschaften, 319. Springer, Berlin, (1999).

[179] M. Talagrand,Transportation cost for Gaussian and other product measures, Geom. Funet. Anal. 6 (1996), 587-600.

[180] M. Talagrand. A new isoperimetric inequality and the concentration of measure phenomenon. In J.Lindenstrouss and V. D. Milman, editors, Geometric Aspects of Functional Analysis, number 1469 in Lecture Notes in Math., pages 94-124, 1991. Springer-Verlag, Berlin.

[181] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. Inst. Hautes Etudes Sci. Pub. Math. 81:73–205, (1995).

[182] M. Talagrand. The supremum of some canonical processes. Amer. J. Math., 116(2):283–325,(1994).

[183] M.A. Rieffel, Metrics on state space, preprint (1999).

[184] M.-K. Renesse, van. Intrinsic coupling on Riemannian manifolds and polyhedra. Elec. J. Probab. 9 (2004), 411-435. Available online at:

[185] M.-K. Von Renesse, Local Poincare via transportation. To appear in Math. Z

[186] M.-K. Von Renesse, & K.-T. Sturm, Transport inequalities, gradient estimates, entropy, -and Ricci curvature. Comm. Pure Appl. Math., 58 (2005), 923-940.

[187] M.L. Eaton, A note on symmetric Bernoulli random variables, Ann. Math. Statist. 41 (1970),1223-1226.

[188] M. Ohya and D. Petz, Quantum Entropy and Use, Sprenger-Verlag, Berlin, 1993

[189] N. Gazlan Poiocare inequalities and dimension free concentration of measure. Preprint, (2008).

[190] N. Gozlan and C. L´eonard. A large deviation approach to some transportation cost inequalities. Probab.Theory Related Fields, 139(1-2):235–283, (2007). [247]

[191] N. Gozlan. Characterization of Talagrand's like transportation-cost inequalities on the real line. J. Funct.Anal., 250(2):400–425, (2007).

[192] N. Gozlan. Integral criteria for transportation cost inequalities. Electron. Comm. Probab., 11:64–77, (2006).

[193] N. J. Korevaar, & R. M. Schoen, Sobolev spaces and harmonic maps for metric space targets. Comm. Anal. Geom., 1 (1993), 561-659.

[194]N.A. Gozlan, A characterization of dimension free concentration in terms of transportation inequalities, preprint, (2008).

[195] O. Rothaus, Diffusion on compact Riemannian manifolds and logarithmic Sobolev inequalities, J. Funct. Anal. 42 (1981), 102-109.

[196] P. A. Deift, Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach, Courant Lecture Notes in Math., Vol. 3, New York Univ., Courant lust. Math. Sei., New York, (1999).

[197] P. Biane and R. speicher, free diffusions, free entropy and free Fisher information, Ann. Inst. H. Poincare probab. Statist. 37 (2001),581-606.

[198] P. Biane, D. Voiculescu, A free probability analogue of the Wasserstein metric on the trace-state space.Geometric and Functional Analysis, 11(6), 1125-1138 (2001).

[199] P. Biane, Free hypercontractivity. Communications in Mathematical Physics, 184, 457-474 (1997). [141]

[200] P. Biane, Logarithmic Sobolev inequalities, matrix models and Free entropy, Acta Math. Sin. (Engl. Ser.) 19 (3) (2003) 497–506. [127]

[201] P. Biane, On the free convolution with a semicircular distribution, Indiana Unv. Math. J. 46 (1997),705-717.

[202] P. Biane, R. Speicher, Free diffusions, free entropy and free Fisher information, preprint DMA, ENS 99-33 (1999):

[203] P. Biane,Processes with free increments, Math. Z. 227 (1998),143-

174.

[204] P. Cattiaux and A. Guillin. Talagrand's like quadratic transportation cost inequalities. J. Math.Pures Appl., 86(9):341-361, (2006).

[205] P. Cattiaux, I. Gentil, and A. Guillin. Weak logarithmic Sobolev inequalities and entropic convergence.Probab. Theory Related Fields,

139(3-4):563–603, (2007).

[206] P. Deift, Orthogonal Ploynomials and Random Matrices; A Riemann-Hilbert Approach, New York University-Courant Institute of Mathematical Sciences- AMS(2000).

[207] P. Federbush, “A partially alternate derivation of a result of Nelson,”J. Math. Phys 10 (1969),50-52.

[208] P. Hajlasz, & P. Koskela, Sobolev meets Poincare. C. R, Acad. Sci. Paris Ser. I Math., 320 :(1995),1211.-1215.

[209] P. Hajlasz, Sobolev spaces on an arbitrary metric space. Potential Anal., 5 (1996), 403-415.

[210] P. Koosis, Introduction to Hp Spaces, Second edition, Cambridge Tracts in Math., Vol. 115, Cambridge . Univ. Press, Cambridge, (1998).

[211] P. Koskela, Upper gradients and poincare inequalities, in Lecture Notes on Analysis in Metric Spaces (Trento, 1999), pp. 55-69. Appunti Corsi Tenuti Docenti Sc. Scuola Norm. Sup., Pisa, 2000.

[212] P. Meyer, Probability and Potentials, Blaisdell Publ. Co., Waltham, Mass., (1966).

[213] P. petersen, Riemannian geometry. Graduate Texts in Mathematics, 171. Springer, New York, 1998.

[214] P.Hajlasz and P.Koskela,Sobolev met Poincare. Mem. Amer. Math. Soc., 145 (2000).

[215] Ph. Blaene and R. Speicher, Free diffusions free entropy and free fisher information, Ann, Inst. H. poincare probab. Statist, 37 (2001), 581-606 .

[216] P.LI, & S.-T. Yau, On the parabolic kernel of the Schrodinger

operator. Acta Math 156 (1986), 153-201.

[217] P.M. Samson. Concentration of measure inequalities for Markov chains and Φ-mixing processes./Ann. Probab., 28(1):416-461, (2000).

[218] R. Blei, L. Ge, Relationships between combinatorial measurements and Orlicz norms, J.Funet. Anal 257(2009) 683-720.

[219] R. Blei, Analysis in integer and fractional dimensions Cambridge university press, Cambridge (2001).

[220] R. Blei, Y. Pores, J.H. Schmorl , fractional products of sets,Random Structures Algothnos 6(1995) 113-119.

[221] R. Holley and D. Stroock, Logarithmic Sobolev inequalities and stochastic Ising models, J. Statist. Phys. 46 (1987),1159-1194.[99]

[222] R, M. Dudley. Real analysis and probability, The Wadsworth& Brooks/Cole Mathematics L Series, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA.(1989).

[223] R. M. Dudley, Real Analysis and Probability. Wadsworth & Brooks/Cole, Pacific Grove, CA, 1989.

[224] R. Abraham, J. E. Marsden, Foundations of mechanics. Second edition. Benjamin/Cummings Publishing Co., Inc., Advanced Book Program, Reading, Mass. (1978).

[225] R. Bhatia, Matrix; Analysis, Sprioger-Verlag, New York, 1996.

[226] R. Blei, Combinatorial dimension and certain norms in harmonic analysis, Amer. J. Math. 106 (1984) 847–887.

[227] R. Blei, T.W. Körner, Combinatorial dimension and random sets, Israel J. Math. 47 (1984) 65–74.

[228] R. C. Blei, Combinatorial dimension: A continuous parameter (to appear in Alta Matematica).

[229] R. C. Blei, Functions of bounded variation and fractional dimension, Proceedings of the Cortona Conference (to appear in Springer Lecture Notes).

[230] R. C. Blei, A uniformity property for A(2) sets and Grothendieck's inequality, Symposia Mathematica, Vol. XXII (1977), 321-336.

[231] R. C. Blei, and T. W. Korner, Combinatorial dimension and random sets (to appear in Israel J. of Mathematics).

[232] R. E. Edwards and K. A. Ross, p-Sidon sets, J. of Functional Analysis, 15 (1974), 404-427. [210]

[233] R. Forman, Bochner's method for cell complexes and combinatorial Ricci curvature, Discrete Comput. Geom, 29 (2003) 323-374.

[234] R. J. McCann, A convexity principle for interacting gases. Adv. Math,128(1997), 153-179.

[235] R. J. McCann, Polar factorization on Riemannian manifolds, preprint, (1999).

[236] R. J. McCann, Existence and uniqueness of monotone measure-preserving maps. Duke Math. J., 80 (1995), 309-323.

[237] R. Jordan, D. Kinderlehrer, & F. OTTO, The variational formulation of the Fokker Planck equation. SIAM J. Math. Anal., 29 (1998), 1-17.

[238] R. Latala and J. O. Wojtaszczyk. On the infimum convolution inequality. Preprint. Availlable on the ArXiv, 2008.

[239] R. Latala and K. Oleszkiewicz. Between Sobolev and Poincare. In geometric aspects of Functional Analysis. Lect. Notes Math., 1745:147-168, (2000).

[240] R. Zimmer, Ergodic theory and semisimple groups, Monographs in Mathematics 81, Birkhauser. Basel (1984).

[241] R.J. Gardner, The Brunn–Minkowski inequality, Bull. Amer. Math. Soc. (N.S.) 39 (3) (2002) 355–405 (in electronic).

[242] R.J. McCann, "Polar factorization of maps on Riemannian manifolds", Geom. Funct. Anal. 11, p. 589-608 (2001).

[243] S. Aida, and D. Stroock, Moment estimates derived from Poincar´e and logarithmic Sobolev inequalities.Math. Res. Lett., 1:75–86, (1994).

[244] S. Aida, T. Masuda, and I. Shigekawa. Logarithmic Sobolev inequalities and exponential integrability. J.Funct. Anal., 126:83–101, (1994).

[245] S. Chevet, Series de variables aleatoires Gaussiennes a valeurs dans $E\hat{\otimes}_{\varepsilon} F$ Application aux produits d'espaces de Wiener abstraits. (preprint) Universite de Clermont-Ferrant. Expose No. XIX.

[246] S. Gal1ot. D. Hulin, Lafontaine, J. Riemannian geometry. 2nd ed. Universitext. Springer, Berlin, (1990).

[247] S. T. Rachev and L. Rüuschendorf, "Mass Transportation Problems," Probability and Its Application, Springer-Verlag, Springer, New York, (1998).

[248] S. Kullback, A lower bound for discrimination information in terms of variation, IEEE, Trans, inform . 4 (1967), 126.-127.

[249] S. Mazur, "Uber schwache Konvergenz in den Raumen (L^P)," Studia Math., 4 (1933), 128-133.

[250] S. Szarek, D. Voiculescu, Volumes of restricted Minkowski sums and the Free analogue of the entropy power inequality, Commun. Math. Phys. 178 (1996) 563-570.

[251] S.-1. OHTA, On measure contraction property of metric measure spaces. Preprint, 2006.

[252] S.G. Bobkov and B. Zegarlinski. Entropy bounds and isoperimetry. Mem. Amer. Math. Soc.,176(829):x+69, (2005).

[253] S.G. Bobkov and C. Houdr´e. Isoperimetric constants for product probability measures. Ann. Probab., 25(1):184–205, (1997).

[254] S.G. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. J. Funct. Anal., 163(1):1–28, (1999).

[255]S.G. Bobkov and M. Ledoux. Poincare's inequalities and Talagrand's concentration phenomenon for the exponential distribution. Probab. Theory Related Fields, 107(3):383–400, (1997).

[256] S.G. Bobkov, I. Gentil, M. Ledoux, Hypercontractivity of Hamilton–Jacobi equations, J. Math. Pures Appl. (9) 80 (7) (2001) 669–696.

[257] S.G. Bobkov and M. Ledoux. From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities. Geom. Funct. Anal., 10(5):1028–1052, (2000).

[258] T. Cabanal-Duvillard, Fluctuations de la loi empirique de grandes matrices aléatoires, Ann. Inst. H. Poincaré Probab.Statist. 37 (3) (2001) 373–402. [284]

[259]T. Cabanal-Duvillard, Probabilités libres et calcul stochastique, application aux grandes matrices aléatoires, Université Paris VI PhD thesis, (1999).

[260] U. Haagerup, Seminar notes on free probability.

[261] V.I. Arnold and B.A. Khesi, "Topological Methods in Hydrodynamics" Springer-Verg, New York, 1998.

[262] V. G. Mazja. Sobolev spaces. Springer Series in Soviet Mathematics. Springer, Berlin, (1985).

[263] W. Beckner. A generalized Poincare inequality for Gaussian measures. Proc. Amer. Math. Sac.,105:397-400, (1989).

[264] W. Gangbo. An elementary proof of the polar factorization of vector-valued functions, Arch. Rational Meeh. Anal. 128 (1994), 381-399.

[265] W. F.Donoghue, Jt., Monotone Matrrix Functions and Anallytic Continuation, Springer-Verlag, Berlin-Heidelberg-New York, (1974).

[266] W. Orlicz, über unbedingte Konvergenz in Funktionenraumen-I, Studia Math., 4 (1933), 33-37.

[267] W. Rudin, Fourier Analysis on Groups, Interscience, New York, (1967).

[268] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions, Comm. Pure Appl. Math. 44 (1991), 375-417.

[269] Y. Burago, M. Gromov, & G. Perelman, A. D. AlekSandrov spaces with curvatures bounded below. Uspekhi Mat. Nauk, 47:2 (1992), 3-51, 222

(Russian) English translation in Russian Math. Surveys, 47:2 (1992), 1-58.

[270] Y. Higuchi, Combinatorial curvature for planar graphs, J. Graph Theory 38 (2001) 220-229.

[271] Y. Ollivier, Ricci curvature of Markov chains on metric spaces, J. Funet. Anal. 256 (2009) 810-864.

[272] Y. Rinott, Multivariate majorization and rearrangement inequalities with some applications to probability and statistics, Israel J. Math. IS (1973), 60-77.

[273] Z. Qian, "Estimates for weighted volumes and applications", Quart. J. Math. Oxford 48, p. 235-242 (1997).

[274];Shawgy Hussein and Mohammed Ali, Rough Curvature bounds and Free Functional Inequalities, phd. Thesis, Sudan University of Science and Technology, Sudan (2014).