Entropic Measure on Multidimensional Spaces

Karl-Theodor Sturm

Abstract. We construct the entropic measure $\mathbb{P}^\beta$ on compact manifolds of any dimension. It is defined as the push forward of the Dirichlet process (another random probability measure, well-known to exist on spaces of any dimension) under the *conjugation map*

$$\mathfrak{C} : \mathcal{P}(M) \to \mathcal{P}(M).$$

This conjugation map is a continuous involution. It can be regarded as the canonical extension to higher dimensional spaces of a map between probability measures on 1-dimensional spaces characterized by the fact that the distribution functions of μ and $\mathfrak{C}(\mu)$ are inverse to each other.

We also present an heuristic interpretation of the entropic measure as

$$d\mathbb{P}^\beta(\mu) = \frac{1}{Z} \exp\left(-\beta \cdot \mathrm{Ent}(\mu|m)\right) \cdot d\mathbb{P}^0(\mu).$$

Mathematics Subject Classification (2000). Primary 99Z99; Secondary 00A00.

Keywords. Class file, journal.

1. Introduction

Gradient flows of entropy-like functionals on the Wasserstein space turned out to be a powerful tool in the study of various dissipative PDEs on Euclidean or Riemannian spaces M, the prominent example being the heat equation. See e.g. the monographs [Vi03, AGS05] for more examples and further references.

In [RS08], von Renesse and the author presented an approach to stochastic perturbation of the gradient flow of the entropy. It is based on the construction of a Dirichlet form

$$\mathcal{E}(u,u) \quad = \quad \int_{\mathcal{P}(M)} \|\nabla u\|^2(\mu) \; d\mathbb{P}^\beta(\mu)$$

where $\|\nabla u\|$ denotes the norm of the gradient in the Wasserstein space $\mathcal{P}(M)$ as introduced by Otto [Ot01]. The fundamental new ingredient was the measure $\mathbb{P}^\beta$

on the Wasserstein space. This so-called *entropic measure* is an interesting and challenging object in its own right. It is formally introduced as

$$d\mathbb{P}^\beta(\mu) = \frac{1}{Z} \exp\left(-\beta \cdot \mathrm{Ent}(\mu|m)\right) \cdot d\mathbb{P}^0(\mu) \tag{1.1}$$

with some (non-existing) 'uniform distribution' $\mathbb{P}^0$ on the Wasserstein space $\mathcal{P}(M)$ and the relative entropy as a potential.
A rigorous construction was presented for 1-dimensional spaces. In the case $M = [0,1]$ it is based on the bijections

$$\mu \quad \overset{(x)=\mu([0,x])}{\longleftrightarrow} \quad f \quad \overset{g=f^{(-1)}}{\longleftrightarrow} \quad g \quad \overset{g(y)=\nu([0,y])}{\longleftrightarrow} \quad \nu$$

between *probability measures, distribution functions* and *inverse distribution functions* (where $f^{(-1)}(y) = \inf\{x \geq 0 : f(x) \geq y\}$ more precisely denotes the 'right inverse' of f). If $\mathfrak{C} : \mathcal{P}(M) \to \mathcal{P}(M)$ denotes the map $\mu \mapsto \nu$ then the entropic measure $\mathbb{P}^\beta$ is just the push forward under $\mathfrak{C}$ of the Dirichlet-Ferguson process $\mathbb{Q}^\beta$. The latter is a random probability measure which is well-defined on every probability space.

For long time it seemed that the previous construction is definitively limited to dimension 1 since it heavily depends on the use of distribution functions (and inverse distribution functions), – objects which do not exist in higher dimensions. The crucial observation to overcome this restriction is to interpret g as the unique *optimal transport map* which pushes forward m (the normalized uniform distribution on M) to μ:

$$\mu = g_* m.$$

Due to Brenier [Br87] and McCann [Mc01] such a 'monotone map' exists for each probability measure μ on a Riemannian manifold of arbitrary dimension. Moreover, also in higher dimensions such a monotone map g has a unique generalized inverse f, again being a monotone map (with generalized inverse being g). This observation allows to define the *conjugation map*

$$\mathfrak{C} : \mathcal{P}(M) \to \mathcal{P}(M), \ \ \mu \mapsto \nu$$

for any compact manifold M. It is a continuous involution. By means of this map we define the entropic measure as follows:

$$\mathbb{P}^\beta := \mathfrak{C}_* \mathbb{Q}^\beta$$

where $\mathbb{Q}^\beta$ denotes the Dirichlet-Ferguson process on M with intensity measure $\beta \cdot m$. (Actually, such a random probability measure exists on every probability space.)

In order to justify our definition of the entropic measure by some heuristic argument let us assume that $\mathbb{P}^\beta$ were given as in (1.1). The identity $\mathbb{Q}^\beta = \mathfrak{C}_* \mathbb{P}^\beta$ then defines a probability measure which satisfies

$$d\mathbb{Q}^\beta(\nu) = \frac{1}{Z} \exp\left(-\beta \cdot \mathrm{Ent}(m|\nu)\right) \cdot d\mathbb{Q}^0(\nu). \tag{1.2}$$

Given a measurable partition $M = \bigcup_{i=1}^{N} M_i$ and approximating arbitrary probability measures ν by measures with constant density on each of the sets M_i of the partition the previous ansatz (1.2) yields – after some manipulations –

$$\begin{aligned}
&\mathbb{Q}^{\beta}_{M_1,\ldots,M_N}(dx) \\
&= \frac{\Gamma(\beta)}{\prod\limits_{i=1}^{N}\Gamma(\beta m(M_i))} \cdot x_1^{\beta\cdot m(M_1)-1}\cdot\ldots\cdot x_{N-1}^{\beta\cdot m(M_{N-1})-1}\cdot x_N^{\beta\cdot m(M_N)-1}\times \\
&\quad\times \delta_{(1-\sum\limits_{i=1}^{N-1} x_i)}(dx_N)dx_{N-1}\ldots dx_1.
\end{aligned}$$

These are, indeed, the finite dimensional distributions of the Dirichlet-Ferguson process.

2. Spaces of Convex Functions and Monotone Maps

Throughout this paper, M will be a compact subset of a complete Riemannian manifold $\hat{M}$ with Riemannian distance d and m will denote a probability measure with support M, absolutely continuous with respect to the volume measure. We assume that it satisfies a Poincaré inequality: $\exists c > 0$

$$\int_M |\nabla u|^2\, dm \geq c \cdot \int_M u^2\, dm$$

for all weakly differentiable $u : M \to \mathbb{R}$ with $\int_M u\, dm = 0$.
For compact Riemannian manifolds, there is a canonical choice for m, namely, the normalized Riemannian volume measure. The freedom to choose m arbitrarily might be of advantage in view of future extensions: For Finsler manifolds and for non-compact Riemannian manifolds there is no such canonical probability measure.

The main ingredient of our construction below will be the Brenier-McCann representation of optimal transport in terms of gradients of convex functions.

Definition 2.1. *A function $\varphi : M \to \mathbb{R}$ is called $d^2/2$-convex if there exists a function $\psi : M \to \mathbb{R}$ such that*

$$\varphi(x) = -\inf_{y\in M}\left[\frac{1}{2}d^2(x,y) + \psi(y)\right]$$

for all $x \in M$. In this case, φ is called generalized Legendre transform *of ψ or* conjugate *of ψ and denoted by*

$$\varphi = \psi^{\mathfrak{c}}.$$

Let us summarize some of the basic facts on $d^2/2$-convex functions. See [Ro70], [Rü96], [Mc01] and [Vi08] for details.[1]

[1]A function φ is $d^2/2$-convex in our sense if and only if the function $-\varphi$ is c-concave in the sense of [Ro70, Rü96, Mc01, Vi08] with cost function $c(x,y) = d^2(x,y)/2$. In our presentation, the $^{\mathfrak{c}}$

Lemma 2.2. (i) *A function* φ *is* $d^2/2$*-convex if and only if*

$$\varphi^{cc} = \varphi.$$

(ii) *Every* $d^2/2$*-convex function is bounded, Lipschitz continuous and differentiable almost everywhere with gradient bounded by* $D = \sup_{x,y\in M} d(x,y)$.

In the sequel, $\mathcal{K} = \mathcal{K}(M)$ will denote the set of $d^2/2$-convex functions on M and $\tilde{\mathcal{K}} = \tilde{\mathcal{K}}(M)$ will denote the set of equivalence classes in $\mathcal{K}$ with $\varphi_1 \sim \varphi_2$ iff $\varphi_1 - \varphi_2$ is constant. $\mathcal{K}$ will be regarded as a subset of the Sobolev space $H^1(M,m)$ with norm

$$\| u \|_{H^1} = \left[\int_M | \nabla u |^2 \, dm + \int_M u^2 dm \right]^{\frac{1}{2}}$$

and $\tilde{\mathcal{K}} = \mathcal{K}/const$ will be regarded as a subset of the space $\tilde{H}^1 = H^1/const$ with norm

$$\| u \|_{\tilde{H}^1} = \left[\int_M | \nabla u |^2 \, dm \right]^{\frac{1}{2}}.$$

Proposition 2.3. *For each Borel map* $g : M \to M$ *the following are equivalent:*

(i) $\exists \varphi \in \tilde{\mathcal{K}} : g = \exp(\nabla\varphi)$ *a.e. on* M*;*
(ii) g *is an optimal transport map from* m *to* $f_* m$ *in the sense that it is a minimizer of* $h \mapsto \int_M d^2(x, h(x)) m(dx)$ *among all Borel maps* $h : M \to M$ *with* $h_* m = g_* m$.

In this case, the function $\varphi \in \tilde{\mathcal{K}}$ in (i) is defined uniquely. Moreover, in (ii) the map f is the unique minimizer of the given minimization problem.
A Borel map $g : M \to M$ satisfying the properties of the previous proposition will be called *monotone map* or *optimal Lebesque transport*. The set of m-equivalence classes of such maps will be denoted by $\mathcal{G} = \mathcal{G}(M)$. Note that $\mathcal{G}(M)$ does *not depend* on the choice of m (as long as m is absolutely continuous with full support)! $\mathcal{G}(M)$ will be regarded as a subset of the space of maps $L^2((M,m)(M,d))$ with metric $d_2(f,g) = \left[\int_M d^2(f(x), g(x)) m(dx)\right]^{\frac{1}{2}}$.
According to our definitions, the map $\Upsilon : \varphi \mapsto \exp(\nabla\varphi)$ defines a bijection between $\tilde{\mathcal{K}}$ and $\mathcal{G}$. Recall that $\mathcal{P} = \mathcal{P}(M)$ denotes the set of probability measures μ on M (equipped with its Borel σ-field).

Proposition 2.4. *The map* $\chi : g \mapsto g_* m$ *defines a bijection between* $\mathcal{G}$ *and* $\mathcal{P}(M)$*. That is, for each* $\mu \in \mathcal{P}$ *there exists a unique* $g \in \mathcal{G}$ *– called* Brenier map *of* μ *– with* $\mu = g_* m$.

The map χ of course strongly depends on the choice of the measure m. (If there is any ambiguity we denote it by χ_m.)

---

stands for 'conjugate'. For the relation between $d^2/2$-convexity and usual convexity on Euclidean space we refer to chapter 4.

Due to the previous observations, there exist canonical bijections Υ and χ between the sets $\tilde{\mathcal{K}}$, $\mathcal{G}$ and $\mathcal{P}$. Actually, these bijections are even homeomorphisms with respect to the natural topologies on these spaces.

Proposition 2.5. *Consider any sequence $\{\varphi_n\}_{n\in\mathbb{N}}$ in $\tilde{\mathcal{K}}$ with corresponding sequences $\{g_n\}_{n\in\mathbb{N}} = \{\Upsilon(\varphi_n)\}_{n\in\mathbb{N}}$ in $\mathcal{G}$ and $\{\mu_n\}_{n\in\mathbb{N}} = \{\chi(g_n)\}_{n\in\mathbb{N}}$ in $\mathcal{P}$ and let $\varphi \in \tilde{\mathcal{K}}$, $g = \Upsilon(\varphi) \in \mathcal{G}$, $\mu = \chi(g) \in \mathcal{P}$. Then the following are equivalent:*

(i) $\varphi_n \longrightarrow \varphi$ *in* $\tilde{H}_1$
(ii) $g_n \longrightarrow g$ *in* $L^2((M,m),(M,d))$
(iii) $g_n \longrightarrow g$ *in m-probability on M*
(iv) $\mu_n \longrightarrow \mu$ *in L^2-Wasserstein distance* d_W
(v) $\mu_n \longrightarrow \mu$ *weakly.*

Proof. $(i) \Leftrightarrow (ii)$ Compactness of M and smoothness of the exponential map imply that there exists $\delta > 0$ such that $\forall x \in M$, $\forall v_1, v_2 \in T_xM$ with $| v_1 |, | v_2 | \leq D$ and $| v_1 - v_2 | < \delta$:

$$\frac{1}{2} \leq d(exp_x v_1, exp_x v_2)/ \mid v_1 - v_2 \mid_{T_xM} \leq 2.$$

Hence, $\varphi_n \longrightarrow \varphi$ in $\tilde{H}^1$, that is $\int_M \mid \nabla\varphi_n(x) - \nabla\varphi(x) \mid^2_{T_xM} m(dx) \longrightarrow 0$, is equivalent to $\int_M d^2(g_n(x), g(x)) m(dx) \longrightarrow 0$, that is, to $g_n \longrightarrow g$ in $L^2((M,m),(M,d))$.
$(ii) \Leftrightarrow (iii)$ Standard fact from integration theory (taking into account that $d(g_n, g)$ is uniformly bounded due to compactness of M).
$(ii) \Leftrightarrow (iv)$ If $\mu_n = (g_n)_*m$ and $\mu_n = g_*m$ then $(g_n, g)_*m$ is a coupling of μ_n and μ. Hence,

$$d_W^2(\mu_n, \mu) \leq \int_M d^2(g_n(x), g(x)) m(dx). \tag{2.1}$$

$(iv) \Leftrightarrow (v)$ Trivial.
$(ii) \Leftrightarrow (iv)$ [Vi08], Corollary 5.21.

□

Remark 2.6. Since M is compact, assertion (ii) of the previous Proposition is equivalent to
(iii') $g_n \longrightarrow g$ in $L^p((M,m),(M,d))$
for any $p \in [1,\infty)$ and similarly, assertion (iv) is equivalent to
(iv') $\mu_n \longrightarrow \mu$ in L^p-Wasserstein distance.

Remark 2.7. In $n = 1$, the inequality in (2.1) is actually an equality. In other words, the map

$$\chi : (\mathcal{G}, d_2) \to (\mathcal{P}, d_W)$$

is an *isometry.* This is no longer true in higher dimensions.

The well-known fact (Prohorov's theorem) that the space of probability measures on a compact space is itself compact, together with the previous continuity results immediately implies compactness of $\tilde{\mathcal{K}}$ and $\mathcal{G}$.

Corollary 2.8. *(i) $\tilde{\mathcal{K}}$ is a compact subset of $\tilde{H}^1$.*
(ii) $\mathcal{G}$ is a compact subset of $L^2((M,m),(M,d))$.

3. The Conjugation Map

Let us recall the definition of the conjugation map $\mathfrak{C}_{\mathcal{K}}:\varphi\mapsto\varphi^{\mathfrak{c}}$ acting on functions $\varphi: M\to\mathbb{R}$ as follows

$$\varphi^{\mathfrak{c}}(x)=-\inf_{y\in M}\left[\frac{1}{2}d^2(x,y)+\varphi(y)\right].$$

The map $\mathfrak{C}_{\mathcal{K}}$ maps $\mathcal{K}$ bijective onto itself with $\mathfrak{C}_{\mathcal{K}}^2=Id$. For each $\lambda\in\mathbb{R}$, $\mathfrak{C}_{\mathcal{K}}(\varphi+\lambda)=\mathfrak{C}_{\mathcal{K}}(\varphi)-\lambda$. Hence, $\mathfrak{C}_{\mathcal{K}}$ extends to a bijection $\mathfrak{C}_{\tilde{\mathcal{K}}}:\tilde{\mathcal{K}}\to\tilde{\mathcal{K}}$. Composing this map with the bijections $\chi:\mathcal{G}\to\mathcal{P}$ and $\Upsilon:\tilde{\mathcal{K}}\to\mathcal{G}$ we obtain involutive bijections

$$\mathfrak{C}_{\mathcal{G}}=\Upsilon\circ\mathfrak{C}_{\tilde{\mathcal{K}}}\circ\Upsilon^{-1}:\mathcal{G}\to\mathcal{G}$$

and

$$\mathfrak{C}_{\mathcal{P}}=\chi\circ\mathfrak{C}_{\mathcal{G}}\circ\chi^{-1}:\mathcal{P}\to\mathcal{P},$$

called *conjugation map* on $\mathcal{G}$ or on $\mathcal{P}$, respectively. Given a monotone map $g\in\mathcal{G}$, the monotone map

$$g^{\mathfrak{c}}:=\mathfrak{C}_{\mathcal{G}}(g)$$

will be called *conjugate map* or *generalized inverse map*; given a probability measure $\mu\in\mathcal{P}$ the probability measure

$$\mu^{\mathfrak{c}}:=\mathfrak{C}_{\mathcal{P}}(\mu)$$

will be called *conjugate measure.*

Example 3.1. *(i) Let $M=S^n$ be the n-dimensional sphere, and m be the normalized Riemannian volume measure. Put*

$$\mu=\lambda\delta_a+(1-\lambda)m$$

for some point $a\in M$ and $\lambda\in\ ]0,1[$. Then

$$\mu^{\mathfrak{c}}=\frac{1}{1-\lambda}1_{M\setminus B_r(a)}\cdot m$$

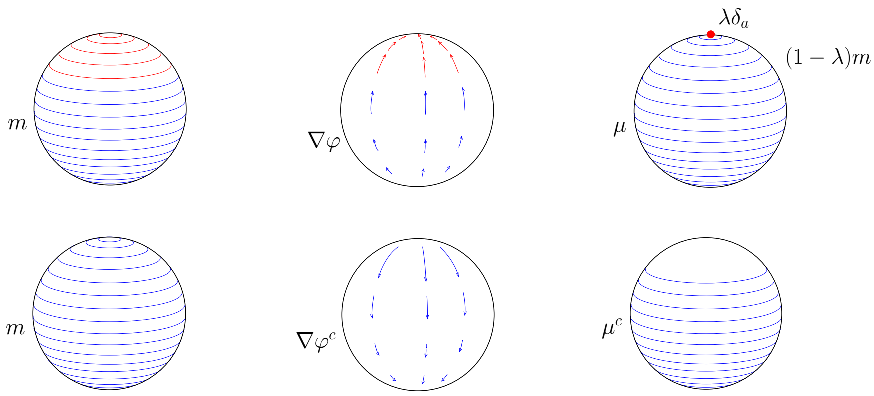
where $r>0$ is such that $m(B_r(a))=\lambda$.

[Proof. The optimal transport map $g=\exp(\nabla\varphi)$ which pushes m to μ is determined by the $d^2/2$-convex function

$$\varphi=\begin{cases}\frac{1}{2}\left[r^2-d^2(a,x)\right] & \text{in } B_r(a)\\ \frac{r}{2(\pi-r)}\left[d^2(a',x)-(\pi-r)^2\right] & \text{in } \overline{B}_{\pi-r}(a')=M\setminus B_r(a)\end{cases}$$

Its conjugate is the function

$$\varphi^{\mathfrak{c}}(y)=-\frac{r}{2\pi}d^2(a',y)+\frac{1}{2}r(\pi-r).\quad]$$



(ii) Let $M = S^n$, the n-dimensional sphere, and $\mu = \delta_a$ for some $a \in M$. Then $\mu^{\mathfrak{c}} = \delta_{a'}$ with $a' \in M$ being the antipodal point of a.

[Proof. Limit of (i) as $\lambda \nearrow 1$. Alternatively: explicit calculations with $\varphi(x) = \frac{1}{2}[\pi^2 - d^2(a,x)]$ and

$$\varphi^{\mathfrak{c}}(y) = \sup_x \left(-\frac{1}{2}d^2(x,y) + \frac{1}{2}d^2(a,x) - \frac{1}{2}\pi^2 \right) = -\frac{1}{2}d^2(a',y). \quad]$$

(iii) Let $M = S^n$, the n-dimensional sphere, and $\mu = \frac{1}{2}\delta_a + \frac{1}{2}\delta_{a'}$ with north and south pole $a, a' \in M$. Then $\mu^{\mathfrak{c}}$ is the uniform distribution on the equator, the $(n-1)$-dimensional set Z of points of equal distance to a, a'.

(iv) Let $M = S^1$ be the circle of length 1, m = uniform distribution and

$$\mu = \sum_{i=1}^{k} \alpha_i \delta_{x_i}$$

with points $x_1 < x_2 < \ldots < x_k < x_1$ in cyclic order on S^1 and numbers $\alpha_i \in [0,1]$, $\sum \alpha_i = 1$. Then

$$\mu^{\mathfrak{c}} = \sum_{i=1}^{k} \beta_i \delta_{y_i}$$

with $\beta_i = \mid x_{i+1} - x_i \mid$ and points $y_1 < y_2 < \ldots < y_k < y_{k+1} = y_1$ on S^1 satisfying

$$\mid y_{i+1} - y_i \mid = \alpha_{i+1}.$$

[Proof. Embedding in $\mathbb{R}^1$ and explicit calculation of distribution and inverse distribution functions.]

Remark 3.2. The conjugation map

$$\mathfrak{C}_{\mathcal{P}} : \mathcal{P} \to \mathcal{P}$$

depends on the choice of the reference measure m on M. Actually, we can choose two different probability measures m_1, m_2 and consider $\mathfrak{C}_{\mathcal{P}} = \chi_{m_2} \circ \mathfrak{C}_{\mathcal{G}} \circ \chi_{m_1}^{-1}$.

Proposition 3.3. *Let $\mu = g_* m \in \mathcal{P}$ be absolutely continuous with density $\eta = \frac{d\mu}{dm}$. Put $f = g^{\mathfrak{c}}$ and $\nu = f_* m = \mu^{\mathfrak{c}}$.*
(i) If $\eta > 0$ a.s. then the measure ν is absolutely continuous with density $\rho = \frac{d\nu}{dm} > 0$ satisfying

$$\eta(x) \cdot \rho(f(x)) = \rho(x) \cdot \eta(g(x)) = 1 \qquad \textit{for a.e. } x \in M.$$

(ii) If ν is absolutely continuous then $f(g(x)) = g(f(x)) = x$ for a.e. $x \in M$.
(iii) Under the previous assumption the Jacobian $\det Df(x)$ and $\det Dg(x)$ exist for almost every $x \in M$ and satisfy

$$\det Df(g(x)) \cdot \det Dg(x) = \det Df(x) \cdot \det Dg(f(x)) = 1,$$

$$\sigma(x) \cdot \eta(x) = \sigma(f(x)) \cdot \det Df(x), \qquad \sigma(x) \cdot \rho(x) = \sigma(g(x)) \cdot \det Dg(x)$$

for almost every $x \in M$ where $\sigma = \frac{dm}{d\mathrm{vol}}$ denotes the density of the reference measure m with respect to the Riemannian volume measure vol.

Proof. (i) For each Borel function $v : M \to \mathbb{R}_+$

$$\int_M v\, d\nu = \int_M v \circ f\, dm = \int_M v \circ f \cdot \frac{1}{\eta}\, d\mu = \int_M v \circ f \cdot \frac{1}{\eta(g \circ f)}\, d\mu = \int_M v \cdot \frac{1}{\eta \circ g}\, dm.$$

Hence, ν is absolutely continuous with respect to m with density $\frac{1}{\eta \circ g}$. Interchanging the roles of μ and ν (as well as f and g) yields the second claim.
(ii), (iii) Part of Brenier- McCann representation result of optimal transports. □

Corollary 3.4. *Under the assumption $\eta > 0$ of the previous Proposition:*

$$\mathrm{Ent}(\mu^{\mathfrak{c}} \mid m) = \mathrm{Ent}(m \mid \mu).$$

Proof. With notations from above

$$\mathrm{Ent}(\mu^{\mathfrak{c}} \mid m) = \int \rho \log \rho\, dm = \int \frac{1}{\eta \circ g} \log \frac{1}{\eta \circ g}\, dm = \int \frac{1}{\eta} \log \frac{1}{\eta}\, d\mu = \mathrm{Ent}(m \mid \mu).$$

□

Lemma 3.5. *The conjugation map*

$$\mathfrak{C}_{\mathcal{K}} : \mathcal{K} \to \mathcal{K}$$

is continuous.

Proof. To simplify notation denote $\mathfrak{C}_{\mathcal{K}}$ by $\mathfrak{C}$. Choose a countable dense set $\{y_i\}_{i \in \mathbb{N}}$ in M and for $k \in \mathbb{N}$ define $\mathfrak{C}_k : \varphi \mapsto \varphi_k^{\mathfrak{c}}$ on $\mathcal{K}$ by $\varphi_k^{\mathfrak{c}}(x) = -\inf_{i=1,\dots,k} [\frac{1}{2} d^2(x, y_i) + (\varphi(y_i)]$. Then as $k \to \infty$

$$\varphi_k^{\mathfrak{c}} \nearrow \varphi^{\mathfrak{c}} \qquad \text{pointwise on M.}$$

Recall that each $\varphi \in \mathcal{K}$ is Lipschitz continuous with Lipschitz constant D.

For each $\varepsilon > 0$ choose $k = k(\varepsilon) \in \mathbb{N}$ such that the set $\{y_i\}_{i=1,\dots,k(\varepsilon)}$ is an ε-covering of the compact space M. Then

$$
\begin{aligned}
&| \mathfrak{C}_k(\varphi)(x) - \mathfrak{C}(\varphi)(x) | \leq \sup_{y\in M} \inf_{i=1,\dots,k} | \frac{1}{2}d^2(x,y) - \frac{1}{2}d^2(x,y_i) + \varphi(y) - \varphi(y_i) | \\
&\leq \sup_{y\in M} \inf_{i=1,\dots,k} 2D \cdot d(y,y_i) \leq 2D\varepsilon \qquad \text{uniformly in } x \in M \text{ and } \varphi \in \mathcal{K}.
\end{aligned}
$$

Now let us consider a sequence $(\varphi_l)_{l\in\mathbb{N}}$ in $\mathcal{K}$ with $\varphi_l \to \varphi$ in $H^1(M)$. Then for each $k \in \mathbb{N}$ as $l \to \infty$

$$\mathfrak{C}_k(\varphi_l) \to \mathfrak{C}_k(\varphi)$$

pointwise on M and thus also in $L^2(M)$. Together with the previous uniform convergence of $\mathfrak{C}_k \to \mathfrak{C}$ it implies

$$\mathfrak{C}(\varphi_l) \to \mathfrak{C}(\varphi)$$

in $L^2(M)$ as $l \to \infty$. Moreover, we know that $\{\mathfrak{C}(\varphi_l)\}_{l\in\mathbb{N}}$ is bounded in $H^1(M)$ (since all gradients are bounded by D). Therefore, finally

$$\mathfrak{C}(\varphi_l) \to \mathfrak{C}(\varphi)$$

in $H^1(M)$ as $l \to \infty$. This proves the continuity of $\mathfrak{C} : \mathcal{K} \to \mathcal{K}$ with respect to the H^1-norm. □

Theorem 3.6. *The conjugation map*

$$\mathfrak{C}_{\mathcal{P}} : \mathcal{P} \to \mathcal{P}$$

is continuous (with respect to the weak topology).

Proof. Let us first prove continuity of the conjugation map $\mathfrak{C}_{\tilde{\mathcal{K}}} : \tilde{\mathcal{K}} \to \tilde{\mathcal{K}}$ (with respect to the $\tilde{H}^1$-norm on $\tilde{\mathcal{K}}$). Indeed, this follows from the previous continuity result together with the facts that the embedding $H^1 \to \tilde{H}^1,\ \varphi \mapsto \tilde{\varphi} = \{\varphi + c : c \in \mathbb{R}\}$ is continuous (trivial fact) and that the map $\tilde{H}^1 \to H^1,\ \tilde{\varphi} = \{\varphi + c : c \in \mathbb{R}\} \mapsto \varphi - \int_M \varphi dm$ is continuous (consequence of Poincaré inequality).
This in turn implies, due to Proposition 2.5, that the conjugation map $\mathfrak{C}_{\mathcal{G}} : \mathcal{G} \to \mathcal{G}$ is continuous (with respect to the L^2-metric on $\mathcal{G}$). Moreover, due to the same Proposition it therefore also implies that the conjugation map

$$\mathfrak{C}_{\mathcal{P}} : \mathcal{P} \to \mathcal{P}$$

is continuous (with respect to the weak topology). □

Remark 3.7. In dimension $n = 1$, the conjugation map $\mathfrak{C}_{\mathcal{G}} : \mathcal{G} \to \mathcal{G}$ is even an isometry from $\mathcal{G}$, equipped with the L^1-metric, into itself.

4. Example: The Conjugation Map on $M \subset \mathbb{R}^n$

In this chapter, we will study in detail the Euclidean case. We assume that M is a compact convex subset of $\mathbb{R}^n$. (The convexity assumption is to simplify notations and results.) The probability measure m is assumed to be absolutely continuous with full support on M.
A function $\varphi : M \to \mathbb{R}$ is $d^2/2$-convex if and only if the function $\varphi_1(x) = \varphi(x) + |x|^2/2$ is *convex* in the usual sense:

$$\varphi_1(\lambda x + (1-\lambda)y) \leq \lambda\varphi_1(x) + (1-\lambda)\varphi_1(y)$$

(for all $x, y \in M$ and $\lambda \in [0,1]$) and if its subdifferential lies in M:

$$\partial\varphi_1(x) \subset M$$

for all $x \in M$.
A function ψ is the conjugate of φ if and only if the function $\psi_1(y) = \psi(y) + |y|^2/2$ is the *Legendre-Fenchel transform* of φ_1:

$$\psi_1(y) = \sup_{x \in M} \left[\langle x, y\rangle - \varphi_1(x)\right].$$

A Borel map $g : M \to M$ is *monotone* if and only if

$$\langle g(x) - g(y), x - y\rangle \geq 0$$

for a.e. $x, y \in M$. Equivalently, g is monotone if and only if $g = \nabla\varphi_1$ for some convex $\varphi_1 : M \to \mathbb{R}$.

Lemma 4.1. *(i) If $\mu = \lambda\delta_z + (1-\lambda)\nu$ then there exists an open convex set $U \subset M$ with $m(U) = \lambda$ such that the optimal transport map g with $g_*m = \mu$ satisfies $g \equiv z$ a.e. on U.*
(ii) The conjugate measure $\mu^{\mathfrak{c}}$ does not charge U:

$$\mu^{\mathfrak{c}}(U) = 0.$$

Proof. (i) Linearity of the problem allows to assume that $z = 0$. Let $g = \nabla\varphi_1$ denote the optimal transport map with φ_1 being an appropriate convex function. Let V be the subset of points in M in which φ_1 is weakly differentiable with vanishing gradient. By the push forward property it follows that $m(V) = \lambda$. Firstly, then convexity of φ_1 implies that φ_1 has to be constant on V, say $\varphi_1 \equiv \alpha$ on V. Secondly, the latter implies that $\varphi_1 \equiv \alpha$ on the convex hull W of V. The interior U of this convex set W has volume $m(U) = m(W) \geq m(V) = \lambda$ and φ_1 is constant on U, hence, differentiable with vanishing gradient. Thus finally $U \subset V$ and $m(U) = \lambda$.
(ii) Let μ_ϵ, $\epsilon \in [0,1]$, denote the intermediate points on the geodesic from $\mu_0 = \mu$ to $\mu_1 = m$. Then $\mu_\epsilon = (g_\epsilon)_*m$ with $g_\epsilon = \exp((1-\epsilon)\nabla\varphi) = \epsilon \cdot Id + (1-\epsilon) \cdot g$ and each μ_ϵ is absolutely continuous w.r. to m. Hence, $g_\epsilon^{\mathfrak{c}} = g_\epsilon^{-1}$ a.e. on M. Therefore, the conjugate measure $\mu_\epsilon^{\mathfrak{c}}$ satisfies

$$\mu_\epsilon^{\mathfrak{c}}(U) = m\left((g_\epsilon^{\mathfrak{c}})^{-1}(U)\right) = m\left(g_\epsilon(U)\right) = \epsilon^n \cdot m(U) = \epsilon^n \cdot \lambda.$$

Now obviously $\mu_\epsilon \to \mu$ as $\epsilon \to 0$. According to Theorem 3.6 this implies $\mu_\epsilon^{\mathfrak{c}} \to \mu^{\mathfrak{c}}$ and thus (since U is open)

$$\mu^{\mathfrak{c}}(U) \le \liminf_{\epsilon \to 0} \mu_\epsilon^{\mathfrak{c}}(U) = 0.$$

□

Theorem 4.2. *(i) If $\mu = \sum_{i=1}^N \lambda_i \delta_{z_i}$ with $N \in \mathbb{N} \cup \{\infty\}$ then there exist disjoint convex open sets $U_i \subset M$ with $m(U_i) = \lambda_i$ such that the optimal transport map $g = \nabla \varphi_1$ with $g_* m = \mu$ satisfies $g \equiv z_i$ on each of the U_i, $i \in \mathbb{N}$.*
The measure $\mu^{\mathfrak{c}}$ is supported by the compact m-zero set $M \setminus \bigcup_{i=1}^N U_i$.
(ii) Each of the sets U_i is the interior of $M \cap A_i$ where

$$A_i = \{x \in \mathbb{R}^n : \varphi_1(x) = \langle z_i, x \rangle + \alpha_i\}$$

and

$$\varphi_1(x) = \sup_{i=1,\dots,N} [\langle z_i, x \rangle + \alpha_i]$$

with numbers α_i to be chosen in such a way that $m(A_i) = \lambda_i$.
(iii) If $N < \infty$ then each of the sets $A_i \subset \mathbb{R}^n$, $i = 1, \dots, N$ is a convex polytope. The decomposition $\mathbb{R}^n = \bigcup_{i=1}^N A_i$ is a Laguerre tesselation (see e.g. [LZ08] *and references therein).*
The compact m-zero set $M \setminus \bigcup_{i=1}^N U_i$ which supports $\mu^{\mathfrak{c}}$ has finite $(n-1)$- dimensional Hausdorff measure.

Corollary 4.3. *(i) If μ is discrete then the topological support of $\mu^{\mathfrak{c}}$ is a m-zero set. In particular, $\mu^{\mathfrak{c}}$ has no absolutely continuous part.*
(ii) If μ has full topological support then $\mu^{\mathfrak{c}}$ has no atoms.

Proof. (i) Obvious from the previous theorem.
(ii) If $\mu^{\mathfrak{c}}$ had an atom (of mass $\lambda > 0$) then according to the previous lemma there would be a convex open set U (of volume $m(U) = \lambda$) such that $\mu(U) = (\mu^{\mathfrak{c}})^{\mathfrak{c}}(U) = 0$. □

5. The Entropic Measure – Heuristics

Our goal is to construct a canonical probability measure $\mathbb{P}^\beta$ on the Wasserstein space $\mathcal{P} = \mathcal{P}(M)$ over a compact Riemannian manifold, according to the formal ansatz

$$\mathbb{P}^\beta(d\mu) = \frac{1}{Z} e^{-\beta \operatorname{Ent}(\mu|m)} \mathbb{P}^0(d\mu).$$

Here $\operatorname{Ent}(\cdot \mid m)$ is the *relative entropy* with respect to the reference measure m, β is a constant > 0 ('the inverse temperature') and $\mathbb{P}^0$ should denote a (non-existing) 'uniform distribution' on $\mathcal{P}(M)$. Z should denote a normalizing constant. Using the conjugation map $\mathfrak{C}_{\mathcal{P}} : \mathcal{P}(M) \to \mathcal{P}(M)$ and denoting $\mathbb{Q}^\beta := (\mathfrak{C}_{\mathcal{P}})_* \mathbb{P}^\beta$, $\mathbb{Q}^0 := (\mathfrak{C}_{\mathcal{P}})_* \mathbb{P}^0$ the above problem can be reformulated as follows:

Construct a probability measure $\mathbb{Q}^\beta$ on $\mathcal{P}(M)$ such that – at least formally –

$$\mathbb{Q}^\beta(d\nu) = \frac{1}{Z} e^{-\beta\,\mathrm{Ent}(m|\nu)} \mathbb{Q}^0(d\nu) \tag{5.1}$$

with some 'uniform distribution' $\mathbb{Q}^0$ in $\mathcal{P}(M)$. Here, we have used the fact that

$$\mathrm{Ent}(\nu^{\mathfrak{c}} \mid m) = \mathrm{Ent}(m \mid \nu)$$

(Corollary 3.4), at least if $\nu \ll m$ with $\frac{d\nu}{dm} > 0$ almost everywhere.
Probability measures $\mathbf{P}(d\mu)$ on $\mathcal{P}(M)$ – so called *random probability measures* on M – are uniquely determined by the distributions $\mathbf{P}_{M_1,\ldots,M_N}$ of the random vectors

$$(\mu(M_1), \ldots, \mu(M_N))$$

for all $N \in \mathbb{N}$ and all measurable partitions $M = \dot\bigcup_{i=1}^N M_i$ of M into disjoint measurable subsets M_i. Conversely, if a consistent familiy $\mathbf{P}_{M_1,\ldots,M_N}$ of probability measures on $[0,1]^N$ (for all $N \in \mathbb{N}$ and all measurable partitions $M = \dot\bigcup_{i=1}^N M_i$) is given then there exists a random probability measure $\mathbf{P}$ such that

$$\mathbf{P}_{M_1,\ldots,M_N}(A) = \mathbf{P}((\mu(M_1), \ldots, \mu(M_N)) \in A)$$

for all measurable $A \subset [0,1]^N$, all $N \in \mathbb{N}$ and all partitions $M = \dot\bigcup_{i=1}^N M_i$.
Given a measurable partition $M = \dot\bigcup_{i=1}^N M_i$ the ansatz (5.1) yields the following characterization of the finite dimensional distribution on $[0,1]^N$

$$\mathbb{Q}^\beta_{M_1,\ldots,M_N}(dx) = \frac{1}{Z_N} e^{-\beta S_{M_1,\ldots,M_N}(x)} q_{M_1,\ldots,M_N}(dx) \tag{5.2}$$

where $S_{M_1,\ldots,M_N}(x)$ denotes the conditional expectation (with respect to $\mathbb{Q}^0$) of $S(\cdot) = \mathrm{Ent}(m \mid \cdot)$ under the condition $\nu(M_1) = x_1, \ldots, \nu(M_N) = x_N$.
Moreover, $q_{M_1,\ldots,M_N}(dx) = \mathbb{Q}^0((\nu(M_1), \ldots, \nu(M_N)) \in dx)$ denotes the distribution of the random vector $(\nu(M_1), \ldots, \nu(M_N))$ in the simplex $\sum_N = \left\{x \in [0,1]^N : \sum_{i=1}^N x_i = 1\right\}$. According to our choice of $\mathbb{Q}^0$, the measure $q_{M_1,\ldots,M_N}$ should be the 'uniform distribution' in the simplex $\sum_N$. In [RS08] we argued that the canonical choice for a 'uniform distribution' in $\sum_N$ is the measure

$$q_N(dx) = c \cdot \frac{dx_1 \ldots dx_{N-1}}{x_1 \cdot x_2 \cdot \ldots \cdot x_{N-1} \cdot x_N} \cdot \delta_{(1-\sum_{i=1}^{N-1} x_i)}(dx_N). \tag{5.3}$$

It remains to get hands on $S_{M_1,\ldots,M_N}(x)$, the *conditional expectation* of $S(\cdot) = \mathrm{Ent}(m \mid \cdot)$ under the constraint $\nu(M_1) = x_1, \ldots, \nu(M_N) = x_N$. We simply replace it by $\underline{S}_{M_1,\ldots,M_N}(x)$, the *minimum* of $\nu \mapsto \mathrm{Ent}(m \mid \nu)$ under the constraint $\nu(M_1) = x_1, \ldots, \nu(M_N) = x_N$.
Obviously, this minimum is attained at a measure with constant density on each of the sets M_i of the partition, that is

$$\nu = \sum_{i=1}^N \frac{x_i}{m(M_i)} 1_{M_i} m.$$

Hence,

$$\underline{S}_{M_1,\ldots,M_N}(x) = -\sum_{i=1}^{N} \log \frac{x_i}{m(M_i)} \cdot m(M_i). \tag{5.4}$$

Replacing $\underline{S}_{M_1,\ldots,M_N}$ by $S_{M_1,\ldots,M_N}$ in (5.2), the latter yields

$$\begin{aligned}
&\mathbb{Q}^{\beta}_{M_1,\ldots,M_N}(dx) = c \cdot e^{-\beta \underline{S}_{M_1,\ldots,M_N}(x)} q_N(dx) \\
&= \frac{\Gamma(\beta)}{\prod_{i=1}^{N} \Gamma(\beta m(M_i))} \cdot x_1^{\beta \cdot m(M_1)-1} \cdot \ldots \cdot x_{N-1}^{\beta \cdot m(M_{N-1})-1} \cdot x_N^{\beta \cdot m(M_N)-1} \times \\
&\quad \times \delta_{(1-\sum_{i=1}^{N-1} x_i)}(dx_N) dx_{N-1} \ldots dx_1.
\end{aligned}$$

This, indeed, defines a projective family! Hence, the random probability measure $\mathbb{Q}^{\beta}$ exists and is uniquely defined. It is the well-known *Dirichlet-Ferguson process*. Therefore, in turn, also the random probability measure $\mathbb{P}^{\beta} = (\mathfrak{C}_{\mathcal{P}})_* \mathbb{Q}^{\beta}$ exists uniquely.

6. The Entropic Measure – Rigorous Definition

Definition 6.1. *Given any compact Riemannian space* (M, d, m) *and any parameter* $\beta > 0$ *the entropic measure*

$$\mathbb{P}^{\beta} := (\mathfrak{C}_{\mathcal{P}})_* \mathbb{Q}^{\beta}$$

is the push forward of the Dirichlet-Ferguson process $\mathbb{Q}^{\beta}$ *(with reference measure* βm*) under the conjugation map* $\mathfrak{C}_{\mathcal{P}} : \mathcal{P}(M) \to \mathcal{P}(M)$.

$\mathbb{P}^{\beta}$ as well as $\mathbb{Q}^{\beta}$ are probability measures on the compact space $\mathcal{P} = \mathcal{P}(M)$ of probability measures on M. Recall the definition of the Dirichlet-Ferguson process $\mathbb{Q}^{\beta}$ [Fe73]: For each measurable partition $M = \dot{\bigcup}_{i=1}^{N} M_i$ the random vector $(\nu(M_1), \ldots, \nu(M_N))$ is distributed according to a Dirichlet distribution with parameters $(\beta\, m(M_1), \ldots, \beta\, m(M_N))$. That is, for any bounded Borel function $u : \mathbb{R}^N \to \mathbb{R}$

$$\begin{aligned}
&\int_{\mathcal{P}(M)} u(\nu(M_1), \ldots, \nu(M_N)) \mathbb{Q}^{\beta}(d\nu) = \\
&\frac{\Gamma(\beta)}{\prod_{i=1}^{N} \Gamma(\beta m(M_i))} \cdot \int_{[0,1]^N} u(x_1, \ldots, x_N) \cdot x_1^{\beta\, m(M_1)-1} \cdot \ldots \cdot x_N^{\beta\, m(M_N)-1} \times \\
&\times \delta_{(1-\sum_{i=1}^{N-1} x_i)}(dx_N) dx_{N-1} \ldots dx_1.
\end{aligned}$$

The latter uniquely characterizes the 'random probability measure' $\mathbb{Q}^{\beta}$. The existence (as a projective limit) is guaranteed by Kolmogorov's theorem.

An alternative, more direct construction is as follows: Let $(x_i)_{i\in\mathbb{N}}$ be an iid sequence of points in M, distributed according to m, and let $(t_i)_{i\in\mathbb{N}}$ be an iid sequence of numbers in $[0,1]$, independent of the previous sequence and distributed according to the Beta distribution with parameters 1 and β, i.e. $\text{Prob}(t_i \in ds) = \beta(1-s)^{\beta-1} \cdot 1_{[0,1]}(s)ds$. Put

$$\lambda_k = t_k \cdot \prod_{i=1}^{k-1}(1-t_i) \qquad \text{and} \qquad \nu = \sum_{k=1}^{\infty} \lambda_k \cdot \delta_{x_k}.$$

Then $\nu \in \mathcal{P}(M)$ is distributed according to $\mathbb{Q}^\beta$ [Se94].
The distribution of ν does not change if one replaces the above 'stick-breaking process' $(\lambda_k)_{k\in\mathbb{N}}$ by the 'Dirichlet-Poisson process' $(\lambda_{(k)})_{k\in\mathbb{N}}$ obtained from it by ordering the entries of the previous one according to their size: $\lambda_{(1)} \geq \lambda_{(2)} \geq \ldots \geq 0$. Alternatively, the Dirichlet-Poisson process can be regarded as the sequence of jumps of a Gamma process with parameter β, ordered according to size.
Note that $m(M_0) = 0$ for a given $M_0 \subset M$ implies that $\nu(M_0) = 0$ for $\mathbb{Q}^\beta$-a.e. $\nu \in \mathcal{P}(M)$. On the other hand, obviously, $\mathbb{Q}^\beta$-a.e. $\nu \in \mathcal{P}(M)$ is discrete. In contrast to that, as a corollary to Theorem 4.3 and in analogy to the 1-dimensional case we obtain:

Corollary 6.2. *If $M \subset \mathbb{R}^n$ then $\mathbb{P}^\beta$-a.e. $\mu \in \mathcal{P}(M)$ has no absolutely continuous part and no atoms. The topological support of $\mu^{\mathfrak{c}}$ is a m-zero set.*

For $\mathbb{P}^\beta$-a.e. $\mu \in \mathcal{P}(M)$ there exist a countable number of open convex sets $U_k \subset M$ ('holes in the support of μ') with sizes $\lambda_k = m(U_k)$, $k \in \mathbb{N}$. The measure μ is supported on the complement of all these holes $M \setminus \bigcup_k U_k$, a compact m-zero set. The sequence $(\lambda_k)_{k\in\mathbb{N}}$ of sizes of the holes is distributed according to the stick breaking process with parameter β. In particular,

$$\mathbb{E}\lambda_k = \frac{1}{\beta}\left(\frac{\beta}{1+\beta}\right)^k \qquad (\forall k \in \mathbb{N}).$$

In average, each hole has size $\leq \frac{1}{1+\beta}$. For large β, the size of the k-th hole decays like $\frac{1}{\beta}\exp(-k/\beta)$ as $k \to \infty$. For small β, $\lambda_{(1)}$ the size of the largest hole is of order $\sim \frac{1}{1+0.7\beta}$, [Gr88].

Remark 6.3. In principle, the reference measures in the conjugation map (see Remark 3.2) and in the Dirichlet-Ferguson process could be chosen different from each other.

Given a diffeomorphism $h : M \to M$ the challenge for the sequel will be to deduce a *change of variable formula* for the entropic measure $\mathbb{P}^\beta(d\mu)$ under the induced transformation

$$\mu \mapsto h_*\mu$$

of $\mathcal{P}(M)$.

Conjecture 6.4. *For each φ^2-diffeomorphism $h : M \to M$ there exists a function $Y_h^\beta : \mathcal{P} \to \mathbb{R}$ such that*

$$\int U(h_*\mu)\mathbb{P}^\beta(d\mu) = \int U(\mu)Y_h^\beta(\mu)\mathbb{P}^\beta(d\mu), \tag{6.1}$$

for all bounded Borel functions $U : \mathcal{P} \to \mathbb{R}$. (It suffices to consider U of the form $U(\mu) = u(\mu(M_1)), \ldots, \mu(M_N))$ for measurable partitions $M = \bigcup M_i$ and bounded measurable $u : \mathbb{R}^N \to \mathbb{R}$.) The density Y_h^β is of the form

$$Y_h^\beta(\mu) = \exp\left(\beta \int_M \log\det Dh(x)\mu(dx)\right) \cdot Y_h^0(\mu) \tag{6.2}$$

with $Y_h^0(\mu)$ being independent of β.

As an intermediate step, in order to derive a more direct representation for the entropic measure $\mathbb{P}^\beta$ on $\mathcal{P}(M)$, we may consider the measure

$$\mathbb{Q}_\mathcal{G}^\beta := (\chi^{-1})_*\mathbb{P}^\beta = (\mathfrak{C}_\mathcal{G} \circ \chi^{-1})_*\mathbb{Q}^\beta$$

on $\mathcal{G}$. It is the unique probability measure on the space $\mathcal{G}$ of monotone maps with the property that

$$\int_\mathcal{G} u(m((g^\mathfrak{c})^{-1}(M_1)), \ldots, m((g^\mathfrak{c})^{-1}(M_N))\mathbb{Q}_\mathcal{G}^\beta(dg) =$$
$$\frac{\Gamma(\beta)}{\prod_{i=1}^N \Gamma(\beta m(M_i))} \cdot \int_{[0,1]^N} u(x_1, \ldots, x_N) \cdot x_1^{\beta\, m(M_1)-1} \cdot \ldots \cdot x_N^{\beta\, m(M_N)-1} \times$$
$$\times\, \delta_{(1-\sum_{i=1}^{N-1} x_i)}(dx_N)dx_{N-1}\ldots dx_1$$

for each measurable partition $M = \dot\bigcup_{i=1}^N M_i$ and each bounded Borel function $u : \mathbb{R}^N \to \mathbb{R}$. Actually, one may assume without restriction that the partition consists of continuity sets of m (i.e. $m(\partial M_i) = 0$ for all $i = 1, \ldots, N$) and that u is continuous. Note that $(g^\mathfrak{c})^{-1} = g$ almost everywhere whenever $g_*m \ll m$. Moreover, note that in dimension 1, say $M = [0,1]$, the map $\mathfrak{C}_\mathcal{G} \circ \chi^{-1} : \mathcal{P} \to \mathcal{G}$ assigns to each probability measure ν its cumulative distribution function g.

In dimension 1, the change of variable formula (6.1) allows to prove closability of the Dirichlet form

$$\mathcal{E}(u,u) \quad = \quad \int_\mathcal{P} \|\nabla u\|^2(\mu)\; d\mathbb{P}^\beta(\mu)$$

and to construct the *Wasserstein diffusion* $(\mu_t)_{t\geq 0}$, the reversible Markov process with continuous trajectories (and invariant distribution $\mathbb{P}^\beta$) associated to it [RS08]. The change of variable formula in dimension 1 can also be regarded as a ‘Girsanov type theorem’ for the (normalized) Gamma process [RYZ07]. Until now, no higher dimensional analogue is known.

The Wasserstein diffusion on 1-dimensional spaces satisfies a logarithmic Sobolev inequality [DS07]; it can be obtained as scaling limit of empirical distributions of interacting particle systems [AR07].

References

[AGS05] L. Ambrosio, N. Gigli and G. Savaré (2005): Gradient flows in metric spaces and in the space of probability measures. *Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel.*

[AR07] S. Andres and M. K. von Renesse(2007): Particle Approximation of the Wasserstein Diffusion. *Submitted.*

[Br87] Y. Brenier (1987): Décomposition polaire et réarrangement monotone des champs de vecteurs. *C. R. Acad. Sci. Paris Sér. I Math.* **305** no. 19, 805–808.

[CMS01] D. Cordero-Erausquin, R.J. McCann and M. Schmuckenschläger (2001): A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. *Invent. Math.* **146** no. 2, 219–257.

[DS07] M. Döring and W. Stannat (2007): The logarithmic Sobolev inequality for the Wasserstein Diffusion. *Preprint.*

[Fe73] T. S. Ferguson (1973): A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1**, 209–230.

[Gr88] R. C. Griffiths (1988): On the Distribution of Points in a Poisson Dirichlet Process. *Journal of Applied Probability* **25** no. 2, 336-345.

[LZ08] C. Lautensack and S. Zuyev (2008): Random Laguerre tessellations. *Adv. in Appl. Probab.* **40** no. 3, 630-650.

[Mc01] R.J. McCann (2001): Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.* **11** no. 3, 589–608.

[Oh08] S. Ohta (2008): Finsler interpolation inequalities. *Preprint.*

[Ot01] F. Otto (2001): The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations* **26** no. 1-2, 101–174.

[Ro70] R.T. Rockafellar (1970): Convex analysis. *Princeton Mathematical Series* **28**.

[RS08] M. K. von Renesse and K. T. Sturm (2008): Entropic measure and Wasserstein Diffusion. *To appear in Ann. Probab.*

[RYZ07] M. K. von Renesse, M.Yor and L. Zambotti (2007): Quasi-invariance properties of a class of subordinators. *To appear in Stoch Proc Appl.*

[Rü96] L. Rüschendorf (1996): On c-optimal random variables. *Statist. Probab. Lett.* **27** no. 3, 267–270.

[Se94] Sethuraman, Jayaram (1994): A constructive definition of Dirichlet priors. *Statist. Sinica* **4** no. 2, 639–650.

[Vi03] C. Villani (2003): Topics in Mass Transportation. *Graduate Studies in Mathematics, American Mathematical Society.*

[Vi08] C. Villani (2008): Optimal Transport - Old and new. *Grundlehren. Springer, Berlin.*

Karl-Theodor Sturm
Institut fuer Angewandte Mathematik
Poppelsdorfer Allee 82
D 53115 Bonn
Germany
e-mail: sturm@uni-bonn.de