A RADEMACHER-TYPE THEOREM
ON $L^2$-WASSERSTEIN SPACES OVER
CLOSED RIEMANNIAN MANIFOLDS∗

BY LORENZO DELLO SCHIAVO†

Universität Bonn

October 25, 2018

Let $P$ be any Borel probability measure on the $L^2$-Wasserstein space $(\mathcal{P}_2(X), W_2)$ over a closed Riemannian manifold $X$. We consider the Dirichlet energy integral $E$ induced by $P$ and by the Wasserstein gradient on $\mathcal{P}_2(X)$. Under natural assumptions on $P$, we show that $W_2$-Lipschitz functions on $\mathcal{P}_2(X)$ are contained in the Dirichlet space $D(E)$ and that $W_2$ is dominated by the intrinsic metric induced by $E$. We illustrate several examples.

Introduction. We consider the $L^2$-Wasserstein space $\mathcal{P}_2 = (\mathcal{P}_2(X), W_2)$ associated to a closed Riemannian manifold $(X, g)$. Since the seminal work of F. Otto [32], the geometry of $\mathcal{P}_2$ has been widely studied from several view points. Definitions have been proposed and thoroughly studied of a ‘weak Riemannian structure’ on $\mathcal{P}_2$ (e.g. Lott [27]), of a gradient for ‘smooth’ functions on $\mathcal{P}_2$, of tangent space to $\mathcal{P}_2$ at a point (See Gigli [21] for a detailed account of several such notions), of an exponential map [21], of a Levi-Civita connection [22], of differential forms [20]. This heuristic picture of $\mathcal{P}_2$ as an infinite-dimensional Riemannian manifold calls for the existence of a measure on $\mathcal{P}_2$ canonically and uniquely associated to the metric structure. As it is the case for a differential manifold, such a measure — if any — would deserve the name of Riemannian volume measure which we shall adopt in the following.

In this framework, the question of the existence of such a Riemannian volume measure on $\mathcal{P}_2$ has been insistently posed (e.g. [7, 21, 34, 37]). In the case of $X = S^1$, M.-K. von Renesse and K.-T. Sturm [34] proposed as a candidate the entropic measure on $\mathcal{P}_2(S^1)$ (Example 4.15). Whereas a suitable definition of entropic measure on $\mathcal{P}_2(X)$ for a closed Riemannian manifold $X$ was given by K.-T. Sturm in [37], most of its properties in this general case remain unknown. Here, we rather address the question of discerning the properties of a volume measure $\mathcal{P}$ on $\mathcal{P}_2$.

By ‘volume measure’ we shall mean any analogue on $\mathcal{P}_2$ of a measure on a differential manifold induced by a volume form via integration.

We do so by proving a Rademacher-type result on the $\mathcal{P}$-a.e. Fréchet differentiability of $W_2$-Lipschitz functions (Thm. 1.4). Namely, we consider a Dirichlet space $\mathcal{F}$ associated to $\mathcal{P}$ and

∗Research supported by the CRC 1060 and the Hausdorff Center for Mathematics (University of Bonn).
†I am grateful to Prof.s K.-T. Sturm, E. W. Lytvynov and N. Gigli for useful remarks and comments and to Prof. A. Eberle for providing the reference [14]. Since part of this research was carried out during the Intense Activity Period on Metric Measure Spaces and Ricci Curvature (September 4–29, 2017), it is a pleasure to thank the Max Planck Institute for Mathematics in Bonn for the hospitality.

MSC 2010 subject classifications: Primary 31C25; secondary 46G99.

Keywords and phrases: Rademacher theorem, Wasserstein spaces, Dirichlet–Ferguson measure, entropic measure, normalized mixed Poisson measures, Malliavin–Shavgulidze measure

arXiv:1810.10227v1 [math.FA] 24 Oct 2018
to a natural gradient, with core the algebra $\mathcal{FC}^\infty$ of cylinder functions induced by smooth potential energies (Dfn. 1.1). Combining the strategy of [35] with the fine analysis of tangent plans performed by N. Gigli in [21], we study, for functions in $\mathcal{F}$, suitable concepts of directional derivative and differential, proving their consistency on $\mathcal{FC}^\infty$. We show that, if $\mathbb{P}$ is quasi-invariant with respect to the family of shifts defining the gradient, then the space of $W_2$-Lipschitz functions is contained in $\mathcal{F}$.

The requirement of the Rademacher property is indeed a natural one for a volume measure. For instance, it was recently shown by G. De Philippis and F. Rindler [10, 1.14] that, if $\mu$ is a positive Radon measure on $\mathbb{R}^d$ such that every Lipschitz function is $\mu$-a.e. differentiable, then $\mu \ll \mathcal{L}^d$. In infinite dimensions, the problem has been addressed in linear spaces (e.g. Bogachev–Mayer-Wolf [5]), in particular on the abstract Wiener space (Enchev–Stroock [15]), and — in the ‘non-flat’, albeit finitary, case — on configuration spaces (Röckner–Schied [35]).

Finally, we detail some examples of measures satisfying, fully or in part, our assumptions. These are mainly taken from the theory of point processes and include normalized mixed Poisson measures, the Dirichlet–Ferguson measure [16], as well as the entropic measure [34] and an image on $\mathcal{P}_2(\mathbb{S}^1)$ of the Malliavin–Shavgulidze measure [29]. We show through these examples how the situation on $\mathcal{P}_2$ is opposite to the aforementioned result in [10]. In particular, there exist mutually singular fully supported measures on $\mathcal{P}_2$ satisfying the Rademacher property.

Auxiliary results are collected in the Appendix, together with a discussion of the notion of ‘tangent bundle’ to $\mathcal{P}_2$ from the point of view of global derivations of the algebra $\mathcal{FC}^\infty$.

CONTENTS

Introduction .............................................................. 1

1 A Rademacher Theorem on $\mathcal{P}_2$ .................................. 3

2 Preliminaries .......................................................... 7
  2.1 Setting and further notation ....................................... 7
  2.2 Lipschitz functions .................................................. 8
  2.3 Dirichlet forms ....................................................... 9
  2.4 Optimal transport ................................................... 10
  2.5 Geometry of $\mathcal{P}_2$ ............................................. 11

3 Proof of the main result .............................................. 13
  3.1 On the differentiability of $W_2$-cone functions .......... 13
  3.2 On the differentiability of functions along flow curves .... 18
  3.3 On the differentiability of Lipschitz functions .......... 20
  3.4 Proof of Theorem 1.4 ............................................. 22

4 Examples .............................................................. 23
  4.1 On assumption (P) .................................................. 23
  4.2 On assumption (B) .................................................. 26
  4.3 Normalized mixed Poisson measures ...................................... 27
  4.4 The Dirichlet–Ferguson measure .................................. 28
  4.5 The entropic measure ............................................. 29
  4.6 An image on $\mathcal{P}$ of the Malliavin–Shavgulidze measure .... 30

5 Appendix ............................................................. 31
  5.1 On the notion of tangent bundle to $\mathcal{P}_2$ ................. 31
  5.2 Auxiliary results on normalized mixed Poisson measures 35
1. A Rademacher Theorem on $\mathcal{P}_2$. Everywhere in the following let $(X, g)$ be a closed (i.e. compact, without boundary) connected smooth $d$-dimensional Riemannian manifold with intrinsic distance $d$ and volume measure $m$.

Let further $\mathcal{P}$ be the space of all Borel probability measures on $X$. Given $\mu_1, \mu_2 \in \mathcal{P}$, we denote by $\text{Cpl}(\mu_1, \mu_2)$ the set of couplings (or transport plans) between $\mu_1$ and $\mu_2$, that is, the set of Borel probability measures on $X \times X$ such that $\text{pr}_X^\ast \pi = \mu_i$ for $i = 1, 2$. In the following, we consider the $L^2$-Wasserstein space $(\mathcal{P}_2, W_2)$ associated to the metric space $(X, d)$. As a consequence of the compactness of $X$, the space $\mathcal{P}_2$ coincides, as a set, with the space $\mathcal{P}$ which we endow with the $L^2$-Wasserstein distance $W_2$. Given measures $\mu, \nu \in \mathcal{P}$, $i = 1, 2$, the latter is defined as

$$W_2(\mu, \nu) := \inf_{\pi \in \text{Cpl}(\mu, \nu)} \left( \int_{X \times X} d\pi(x, y) d^2(x, y) \right)^{1/2}.$$ 

We denote by $\text{Opt}(\mu, \nu)$ the set of optimal plans $\pi \in \text{Cpl}(\mu, \nu)$ attaining the infimum in (1.1). This set is always non-empty. It is well-known (see e.g. [2] or [42, Chapter 6]) that, under our assumptions on $X$, the space $(\mathcal{P}_2, W_2)$ is a compact (in particular: complete and separable) geodesic metric space.

In order to perform computations for functions on $\mathcal{P}$ in the spirit of [27, 32], we recall the definition of potential energy — in the sense of [41, §5.2.2]. Namely, given a continuous function $f : X \to \mathbb{R}$, we define the potential energy $f^{**} : \mathcal{P} \to \mathbb{R}$ associated to $f$ by setting

$$f^{**} \mu := \mu f = \int_X d\mu f.$$

The notation $f^{**}$ is motivated by a functional analysis perspective: by $f^{**}$ we mean the image of $f$ under the canonical injection of the space of continuous functions $C(X)$ into its bidual.

**Definition 1.1 (Cylinder functions).** For $f_i \in C(X), i \leq k$, we set $f := (f_1, \ldots, f_k)$ and $f^{**} : \mathcal{P} \ni \mu \mapsto (f^{**} \mu, \ldots, f_k^{**} \mu) \in \mathbb{R}^k$, and define the algebra of cylinder functions on $\mathcal{P}$

$$\mathcal{FC}^\infty := \left\{ u : \mathcal{P} \to \mathbb{R} \mid u = F \circ f^{**} \text{ for some } k \in \mathbb{N}, F \in C^\infty(\mathbb{R}^k), f_i \in C^\infty(X) \right\}.$$ 

**Remark 1.2.** By compactness of $\mathcal{P}_2$, in the definition above one might equivalently take $F \in C_c^\infty(\mathbb{R}^k)$. The given definition makes more apparent that $f^{**} \in \mathcal{FC}^\infty$ for all $f \in C^\infty(X)$. By continuity of $f^{**}$, cylinder functions are continuous and thus (Borel) measurable.

Motivated by the analogous choice in the framework of configuration spaces (cf. [35, (1.1)], see §4.3 below), we define the gradient of $u \in \mathcal{FC}^\infty$ by

$$\nabla u(\mu)(x) := \sum_{i=1}^k (\partial_i F)(f^{**} \mu) \nabla f_i(x).$$ 

This choice is consistent, by chain rule, with the Fréchet differentiability of $f^{**}$ with respect to a natural Riemannian structure on the space of absolutely continuous measures $\mu = \rho m \in \mathcal{P}$ (cf.
e.g. [27] or [41, §9.1]) and more generally with the differentiability of functionals on probability measures (e.g. [3]); furthermore, it is also consistent with the definition of a Wasserstein gradient in the recent work [8] (see in particular [8, 2.3 and 2.4]).

We will also need a concept of directional derivative for functions in \( FC^\infty \) and thus a concept of ‘direction’ at a point \( \mu \in \mathcal{P} \). It is not surprising that such a definition ought to be “inherited” from the differential structure of the manifold \( X \), henceforth the base space. Indeed, let \( T_x X \) be the tangent space to \( X \) at the point \( x \). We denote by \( \mathfrak{X}^m \) the space of \( m \)-differentiable vector fields, that is, sections of the tangent bundle \( TX \), endowed with the usual \( C^m \)-norm \( \| \cdot \|_{\mathfrak{X}^m} \). For any \( w \in \mathfrak{X}^\infty \) we denote by \( (\psi^{w,t})_{t \in \mathbb{R}} \) the flow generated by \( w \), i.e. a map \( \psi^{w,t} : X \to X \) such that

\[
\forall x \in X \quad \dot{\psi}^{w,t}(x) = w(\psi^{w,t}(x)) \quad \text{and} \quad \psi^{w,0}(x) = x,
\]

where by \( \dot{\psi}^{w,t}(x) \) we mean the velocity of the curve \( s \mapsto \psi^{w,s}(x) \) at time \( t \). By compactness of \( X \) every \( w \in \mathfrak{X}^\infty \) admits a unique flow, well-defined and a smooth orientation-preserving diffeomorphism in \( \text{Diff}_+(X) \) for all times \( t \in \mathbb{R} \). (See e.g. [4, §1.3.7(ii)].) If we denote by

\[
\Psi^{w,t} := \psi^{w,t} : \mathcal{P} \to \mathcal{P}
\]

the push-forward via \( \psi^{w,t} \), then a straightforward computation (see Lem. 5.2 below) shows that

\[
(\nabla_w u)(\mu) := d_{|t=0}(u \circ \Psi^{w,t})(\mu) = \langle \nabla u(\mu) \mid w \rangle_{\mathfrak{X}_\mu}, \quad u \in FC^\infty,
\]

where, for vector fields \( w^i \in \mathfrak{X}^\infty, i = 0, 1 \), we set

\[
\langle w^0 \mid w^1 \rangle_{\mathfrak{X}_\mu} := \int_X d\mu(x) \langle w^0_x \mid w^1_x \rangle_g.
\]

This would motivate (cf. [35] for the case of configuration spaces) to define the tangent space to \( \mathcal{P} \) at a point \( \mu \) as the space \( \mathfrak{X}_\mu := \text{co}_\mu \mathfrak{X}^\infty \), that is, the abstract linear completion of \( \mathfrak{X}^\infty \) with respect to the norm \( \| \cdot \|_{\mathfrak{X}_\mu} \) induced by the pre-Hilbert scalar product \( \langle \cdot \mid \cdot \rangle_{\mathfrak{X}_\mu} \). We shall also write \( T^\text{Der}_\mu \mathcal{P}_2 \) for \( \mathfrak{X}_\mu \) and thus \( T^\text{Der} \mathcal{P}_2 \) for the associated fiber-"bundle". In the optimal transport literature however (e.g. [2, 20, 21, 22]), it is well-established that one should define instead \( T^\text{V} \mathcal{P}_2 := \text{cl}_{\mathfrak{X}_\mu} \mathfrak{X}^\infty \), where \( \mathfrak{X}^\infty := \nabla C^\infty(X) \) denotes the family of vector fields of gradient type; the associated fiber-"bundle" will be denoted by \( T^\text{V} \mathcal{P}_2 \). In the following we will make use of both non-equivalent\(^1\) definitions. An exhaustive discussion of this choice is postponed to §5.1.

We consider the class of Borel probability measures on \( \mathcal{P}_2 \) satisfying

**Assumption (P).** We say that \( \mathbb{P} \) satisfies (P) if and only if each of the following holds:

\begin{enumerate}[(P)_1]
    \item \( \mathbb{P} \) is fully supported;
    \item \( \mathbb{P} \) is diffuse (i.e. it has no atoms);
    \item \( \mathbb{P} \) satisfies the following integration by parts formula. If \( u, v \in FC^\infty \) and \( w \in \mathfrak{X}^\infty \), then there exists a measurable function \( \mu \mapsto \nabla_w v \in \mathfrak{X}_\mu \) such that
        \[
        \int_\mathcal{P} d\mathbb{P} \nabla_w u \cdot v = \int_\mathcal{P} d\mathbb{P} u \cdot \nabla_w v;
        \]
\end{enumerate}

\(^1\)It is to be noted that the two definitions are however equivalent on configuration spaces.
(P₄) \( P \) is quasi-invariant with respect to the action of the family of flows \( \text{Flow}(X) \) on \( P \), i.e. \( P \) and \( \Psi^{w,t}_t P \) are mutually absolutely continuous for all \( w \in X^\infty \) and \( t \in \mathbb{R} \). Moreover, for all finite \( s \leq t \) it holds that

\[
\text{essinf}_{r \in [s,t]} R^w_r(\mu) > 0 \quad \text{where} \quad R^w_r := \frac{d(\Psi^{w,t}_r P) \otimes dr}{dP \otimes dr}.
\]

The validity and necessity of these assumptions are widely illustrated through examples in §4.

**Definition 1.3 (Cylinder vector fields).** Let \( \mathcal{XC}^\infty := \mathcal{FC}^\infty \otimes \mathbb{R} X^\infty \) denote the vector space of cylinder vector fields on \( P \), i.e. the \( \mathbb{R} \)-vector space of sections \( W \) of \( T \text{Der} P \) of the form

\[
W(\mu)(x) = \sum_j^n v_j(\mu) w_j(x)
\]

with \( n \in \mathbb{N} \), \( v_j \in \mathcal{FC}^\infty \) and \( w_j \in X^\infty \). By \( \mathcal{XC}_P \) we mean the abstract linear completion of the space \( \mathcal{XC}^\infty \) endowed with the pre-Hilbert norm defined by setting

\[
\| W \|^2_{\mathcal{XC}_P} := \sum_j^n \int_P dP(\mu) |v_j(\mu)|^2 \| w_j \|^2_{\mathcal{XC}_P}.
\]

It follows by linearity from assumption \((P_3)\) that

\[
\forall u \in \mathcal{FC}^\infty \quad \forall W \in \mathcal{XC}^\infty \quad \int_P dP \langle \nabla u \mid W \rangle_X = -\int_P dP u \text{div}_P W
\]

where, for any \( W \) as in (1.7),

\[
\text{div}_P W(\mu) := -\sum_i^n \nabla_{w_i} v_i(\mu).
\]

Then, \((\text{div}_P, \mathcal{XC}^\infty)\) is a densely defined linear operator from the space of sections \( \Gamma_{L^2} T\text{Der} P \) to \( L^2(\mathcal{P}) \) and we denote its adjoint by \((d_P, W^{1,2})\). By definition, functions in \( W^{1,2} \) are weakly differentiable, in the sense that (1.8) holds for all \( u \in W^{1,2} \) with \( dP u \) in lieu of \( \nabla u \).

We denote by \( \mathcal{F} \) the set of all bounded measurable functions \( u \) on \( P \) for which there exists a measurable section \( D u \) of \( T\text{Der} P \) such that

\[
\mathcal{E}(u, u) := \int_P dP(\mu) \langle Du(\mu) \mid Du(\mu) \rangle_{X_\mu} < \infty
\]

and such that for every \( w \in X^\infty \) and \( s \in \mathbb{R} \) there exists the directional derivative

\[
L^2(\mathcal{P}, \Psi^{w,s}_s P) - \lim_{t \to 0} \frac{u \circ \Psi^{w,t}_t - u}{t} = \langle Du \mid w \rangle_X.
\]

Finally, set \( \mathcal{F}_{\text{cont}} := \mathcal{F} \cap \mathcal{C}(\mathcal{P}) \) and observe that \( \mathcal{FC}^\infty \subset \mathcal{F}_{\text{cont}} \subset \mathcal{F} \) and that, a priori, every inclusion may be a strict one.

Before stating the main result, we introduce the following — quite restrictive — assumption on the base space. We will comment extensively about this assumption, and about its connection with the Ma–Trudinger–Wang curvature condition, in §4.2.

\[\text{Here, by } \otimes_\mathbb{R} \text{ we mean the algebraic } \mathbb{R}-\text{tensor product.}\]
Assumption (B). We say that $X$ satisfies assumption (B) if, whenever $\mu, \nu \in \mathcal{P}$, $\mu, \nu \ll m$ with smooth nowhere vanishing densities, then there exists a smooth optimal transport map $g$ mapping $\mu$ to $\nu$ (in the sense of Thm. 2.8 below).

**Theorem 1.4.** Suppose that $\mathbb{P}$ satisfies assumptions (P$_2$) and (P$_3$). Then,

1. the bilinear forms $(\mathcal{E}, \mathcal{F}^\infty)$, $(\mathcal{E}, \mathcal{F}_{cont})$ and $(\mathcal{E}, \mathcal{F})$ are closable and their closures, respectively denoted by $(\mathcal{E}, \mathcal{F}_0)$, $(\mathcal{E}, \mathcal{F}_{cont})$ and $(\mathcal{E}, \mathcal{F})$ are strongly local Dirichlet forms.
2. for each $u \in \mathcal{F}$ there exists a measurable section $D_u$ of the tangent bundle $T^{Der} \mathcal{P}_2$ such that

\[
(1.10) \quad D_u = \nabla u, \quad u \in \mathcal{F}^\infty,
\]
and

\[
(1.11) \quad \mathcal{E}(u, u) = \int_\mathbb{P} d\mathbb{P}(\mu) \|D_u(\mu)\|_{X_\mu}^2,
\]

i.e. the form $(\mathcal{E}, \mathcal{F})$ admits carré du champ $\Gamma(u)(\mu) := \|D_u(\mu)\|_{X_\mu}^2$.

3. (Rademacher property) let $u : \mathcal{P} \to \mathbb{R}$ be $W_2$-Lipschitz continuous. Then $u \in \mathcal{F}_{cont}$ and, if additionally (B) holds, then $u \in \mathcal{F}_0$. Furthermore, there exist a measurable set $\Omega^u \subset \mathcal{P}$ of full $\mathbb{P}$-measure and a measurable section $D_u$ of $T^{Der} \mathcal{P}_2$, satisfying (1.10) and (1.11), such that

3.i) for all $\mu \in \Omega^u$ it holds that $\|D_u(\mu)\|_{X_\mu} \leq \text{Lip}[u]$;
3.ii) if additionally (P$_4$) holds, then

\[
(1.12) \quad \forall \mu \in \Omega^u \quad \|D_u(\mu)\|_{X_\mu} \leq |D_u|(\mu),
\]

where $|D_u|$ is the slope of $u$ (see (2.2) below), and, for all $w \in X^\infty$

\[
(1.13) \quad \lim_{t \to 0} \left( \frac{u \circ \Psi_{w,t} - u}{t} \right)(\cdot) = (D_u(\cdot) | w)_{X},
\]

pointwise on $\Omega^u$ and in $L^2(\mathbb{P})$.

We now collect some remarks on the statement of our main theorem.

**Remark 1.5.** As already noticed in the case of configuration spaces (cf. [35, 1.4(iii)]), the Dirichlet forms $(\mathcal{E}, \mathcal{F}_0)$, $(\mathcal{E}, \mathcal{F}_{cont})$ and $(\mathcal{E}, \mathcal{F})$ do in principle differ. A sufficient condition for their coincidence is the essential self-adjointness of the generator of $(\mathcal{E}, \mathcal{F})$ on the core $\mathcal{F}^\infty$.

**Remark 1.6.** It is readily seen that, by compactness of $\mathcal{P}_2$ and the Stone–Weierstraß Theorem, the spaces $\mathcal{F}^\infty$ and $\mathcal{F}_{cont}$ are uniformly dense in $\mathcal{C}(\mathcal{P}_2)$. Together with the Theorem, this implies that the Dirichlet forms $(\mathcal{E}, \mathcal{F}_0)$ and $(\mathcal{E}, \mathcal{F}_{cont})$ are regular strongly local Dirichlet forms on $\mathcal{P}_2$.

**Remark 1.7** (On the definition of ‘volume measure’ on $\mathcal{P}_2$). Assumptions (P$_1$) and (P$_2$) are of a general kind, whereas assumptions (P$_3$) and (P$_4$) are — as already noticed in [35, Rmk. p. 329] for measures on configuration spaces — specifically proper of a volume measure
(as discussed in the Introduction). In particular, assumption \( (P_3) \) may be regarded as a form of ‘gradient-divergence duality’ for \( \mathcal{P} \). Assumption \( (P_4) \) (and its stronger version \( (P_5) \); see §4.1 below) is also expected from a differential geometry point of view and it is equally important in light of Proposition 4.6 below.

**Remark 1.8** (On the definition of ‘Rademacher-type’ properties). Assume we have already shown that \( u_n : \mu \mapsto W_2(\nu, \mu) \) belongs to \( \mathcal{F}_{\text{cont}} \), resp. \( \mathcal{F}_0 \), (cf. Lem.s 3.3 and 3.4 below) and \( \Gamma(u_n) \leq 1 \). Then, \( (3.1) \) may be deduced by the general results on (non-local) Dirichlet forms in [18]. On the contrary — even if it is proven that the Dirichlet form \((\mathcal{E}, \mathcal{F})\) is strongly local and regular — the finer estimate \((1.12)\) does not follow by [24, 2.1], where the reference measure (in our case \( \mathcal{P} \)) is assumed to be doubling. In fact it may be proved that no (fully supported) doubling measure exists on \( \mathcal{P}_2 \), since the latter is infinite-dimensional.

Both of the previous results may be considered as ‘Rademacher-type’ properties for the Dirichlet form(s) in question. Nonetheless, in the case of the Wasserstein space \( \mathcal{P}_2 \), we have — in addition to the general assumptions of [18] or [24] — a good notion of directional derivative for functions on \( \mathcal{P}_2 \). As a consequence, the statement of what we call a ‘Rademacher Theorem on \((\mathcal{P}_2, W_2, \mathcal{P})\)’ comprises more properly assertion \((3.ii)\), where we check that each directional derivative of a “differentiable” function \( u \in \mathcal{F} \) along a “smooth direction” \( w \in X^\infty \) coincides with the scalar product of the “gradient” \( Du \) and “direction” \( w \).

To conclude this preliminary section we anticipate that the statement of our main theorem is non-void, and that our assumptions pose no restriction to the subset of measures in \( \mathcal{P} \) whereon \( \mathcal{P} \) is concentrated. In particular, we prove

**Theorem** (See Rmk. 4.19). Define

- \( \mathcal{A}_1 \) the set of measures in \( \mathcal{P} \) absolutely continuous w.r.t. the volume of \( X \);
- \( \mathcal{A}_2 \) the set of measures in \( \mathcal{P} \) singular continuous w.r.t. the volume of \( X \);
- \( \mathcal{A}_3 \) the set of purely atomic measures in \( \mathcal{P} \);
- \( \mathcal{A}_4 \) the set of transport-regular measures in \( \mathcal{P} \) (see Def. 2.6 below).

Then, \( X = S^1 \) satisfies assumption \((B)\), and, for any choice of \( a_1, a_2, a_3 \geq 0 \) and such that \( a_1 + a_2 + a_3 = 1 \), there exists \( \mathcal{P} \in \mathcal{P}(\mathcal{P}_2) \), satisfying assumption \((P)\) and such that \( \mathcal{P}(\mathcal{A}_i) = a_i \) for every \( i = 1, 2, 3 \) and \( \mathcal{P}(\mathcal{A}_4) = a_1 + a_2 \).

2. Preliminaries.

2.1. Setting and further notation. By a measure we always mean a non-negative measure. We denote by \( I \), resp. \( I^0 \), the unit interval \([0,1]\), resp. \((0,1)\), always endowed with the usual metric, \( \sigma \)-algebra and with the one-dimensional Lebesgue measure \( d\mathcal{L}^1(r) = dr \).

**Measure theoretical setting.** Everywhere in the following let \((Y, \tau)\) be any second countable locally compact Hausdorff topological space with Borel \( \sigma \)-algebra \( \mathcal{B} \) and let \( \mathfrak{n} \) be a \( \sigma \)-finite or (totally) finite fully supported Radon measure on \((Y, \mathcal{B})\). As it is well-known, each and every such \((Y, \tau)\) is a locally compact Polish space (that is, it is separable and completely metrizable), and any finite measure on \((Y, \mathcal{B})\) is a Radon measure. Recall that for any \( \mathcal{B} \)-measurable real-valued function \( f : Y \to \mathbb{R} \), the measure \( |f| \mathfrak{n} \) has a unique (closed) support and set \( \text{supp}[f] := \text{supp}(|f| \mathfrak{n}) \).

If \( f \) is continuous, then \( \text{supp}[f] \) is independent of \( \mathfrak{n} \) in the class of fully supported measures on \((Y, \mathcal{B})\) and it coincides with \( \text{supp} f := \text{cl}_r \{ y \mid f(y) \neq 0 \} \).


**Probability measures on X.** We indicate by $\mathcal{P}^m \subset \mathcal{P}$ the space of probability measures $\mu \ll \mathcal{P}$, by $\mathcal{P}^\infty$ the subset of probability measures $\mu \in \mathcal{P}^m$ with smooth densities, by $\mathcal{P}^\infty \times$ the subset of measures in $\mathcal{P}^\infty$ whose densities with respect to $\mathcal{P}$ are bounded away from 0 (the boundedness (from above) of such densities is rather a consequence of their continuity and of the compactness of $X$). We denote further by $\eta$ any purely atomic measure in $\mathcal{P}$. Usually, we think of any such $\eta$ as an infinite marked configuration and thus we write, with slight abuse of notation, $\eta_x$ in place of $\eta\{x\}$ and $x \in \eta$ whenever $\eta_x > 0$. We denote further by ptws $\eta$ the set of points $x \in X$ such that $\eta_x > 0$, termed the pointwise support of $\eta$. For $r$ in $I$ and any $\mu \in \mathcal{P}$ also set

$$
(2.1) \quad \mu + r \delta_x := (1 - r)\mu + r\delta_x.
$$

2.2. Lipschitz functions. Everywhere in this section let $\rho$ be any metric metrising $(Y, \tau)$. We say that a real-valued function $h : Y \rightarrow \mathbb{R}$ is L-Lipschitz (with respect to $\rho$) if there exists a constant $L > 0$ such that

$$
\forall y_1, y_2 \in Y \quad |h(y_1) - h(y_2)| \leq L \rho(y_1, y_2),
$$

in which case we denote by $\text{Lip}_\rho[h]$ the infimal such constant and by

$$
(2.2) \quad |Dh|_\rho(y) := \limsup_{z \to y} \frac{|h(y) - h(z)|}{\rho(y, z)} \leq L
$$

the slope (or local Lipschitz constant) of $h$ at a point $y \in Y$. The metric $\rho$ is omitted in the notation whenever apparent from context. We set $\rho_\varepsilon(\cdot, \cdot) := \rho(\cdot, \cdot)$ and, for any $A := (a_i)_i^\varepsilon \subset \mathbb{R}$ and $E := (z_i)_i^\varepsilon \subset Y$, we let $\rho_{A, E, L}(\cdot) := \land_{i \leq n}(a_i - L \rho_\varepsilon(\cdot))$.

**Lemma 2.1.** Assume that $(Y, \tau)$ is additionnaly compact, let $Z \subset Y$ be a dense set and fix $h \in \text{Lip}_\rho Y$. For $\varepsilon > 0$ let further $E_\varepsilon := (z_{\varepsilon,i})_i^\varepsilon \subset Z$ be an $\varepsilon$-net\(^3\) for $Y$ and set $A_\varepsilon := (h(z_{\varepsilon, i}))_i^\varepsilon \subset \mathbb{R}$. Then, the function $h_\varepsilon := \rho_{A_\varepsilon, E_\varepsilon, \text{Lip}[h]}$ satisfies $\text{Lip}[h_\varepsilon] \leq \text{Lip}[h]$ and $\|h - h_\varepsilon\|_{C_0} \leq C\varepsilon$ where $C$ is a constant only depending on $\text{Lip}[h]$.

**Proof.** The function $h_\varepsilon$ is $\rho$-Lipschitz continuous with $\text{Lip}[h_\varepsilon] \leq \text{Lip}[h]$ for it is a maximum of $\rho$-Lipschitz continuous functions with Lipschitz constant $\text{Lip}[h]$. Since $h$ is Lipschitz continuous, it coincides with its lower McShane extension [30], i.e. $h(y) = \sup_{z \in Y} \{h(z) - \rho(y, z)\}$. Thus, $h_\varepsilon \leq h$. Furthermore, for all $y \in Y$ there exists $\tilde{z} := \tilde{z}(y)$ such that $h(y) \leq h(\tilde{z}) - \rho(y, \tilde{z}) + \varepsilon$ and, by definition of $E_\varepsilon$, there exists $\tilde{1} := \tilde{1}(y)$ such that $\rho(\tilde{z}, z_{\varepsilon,i}) \leq \varepsilon$. Hence,

$$
\begin{align*}
    h_\varepsilon(y) &\leq h(y) \leq h(\tilde{z}) - \rho(y, \tilde{z}) + \varepsilon \\
    &\leq h(\tilde{z}) - h(z_{\varepsilon,i}) + h(z_{\varepsilon,i}) - \rho(y, z_{\varepsilon,i}) - \rho(y, z_{\varepsilon,i}) + \varepsilon \\
    &\leq h(z_{\varepsilon,i}) - \rho(y, z_{\varepsilon,i}) + h(\tilde{z}) - h(z_{\varepsilon,i}) + |\rho(y, z_{\varepsilon,i}) - \rho(y, \tilde{z})| + \varepsilon \\
    &\leq h_\varepsilon(y) + \text{Lip}[h]\varepsilon + \varepsilon + \varepsilon
\end{align*}
$$

respectively by definition of $h_\varepsilon$, Lipschitz continuity of $h$ and by reverse triangle inequality and definition of $z_{\varepsilon,i}$. The conclusion follows by letting $C := \text{Lip}[h] + 2$.

---

\(^3\)That is, $E_\varepsilon$ is such that $\rho(z_{\varepsilon,i}, z_{\varepsilon,j}) > \varepsilon/2$ for all $i \neq j$ and $\sup_{y \in Y} \rho(y, E_\varepsilon) \leq \varepsilon$. The existence of such an $\varepsilon$-net follows by density of $Z$ in $Y$ and compactness of $Y$. 

\[\Box\]
2.3. Dirichlet forms. We recall some facts on Dirichlet forms and prove some auxiliary results. Whenever \((Q, \mathcal{D}(Q))\) is a non-negative definite symmetric bilinear form, we denote by the same symbol the associated quadratic form, defined as \(Q(u) := Q(u, u)\) if \(u \in \mathcal{D}(Q)\) and \(Q(u) := +\infty\) otherwise.

**Definition 2.2** (Energy measure, carré du champ, intrinsic distance). Let \((\mathcal{E}, \mathcal{D}(\mathcal{E}))\) be a regular strongly local\(^4\) Dirichlet form on \(L^2_n(Y)\). Then (see e.g. [6]), the form \(\mathcal{E}\) can be written as

\[\mathcal{E}(u, v) = \int_Y d\Gamma(u, v)\]

for all \(u, v \in \mathcal{D}(\mathcal{E})\), where \(\Gamma\), termed the energy measure of \((\mathcal{E}, \mathcal{D}(\mathcal{E}))\), is an \(\mathcal{M}(Y, \mathcal{B})\)-valued non-negative definite symmetric bilinear form defined by the formula

\[\int_Y \phi d\Gamma(u, v) := \frac{1}{2} (\mathcal{E}(u, \phi v) + \mathcal{E}(v, \phi u) - \mathcal{E}(uv, \phi))\]

for all \(u, v \in \mathcal{D}(\Gamma) := \mathcal{D}(\mathcal{E}) \cap L^\infty(Y)\) and \(\phi \in \mathcal{D}(\mathcal{E}) \cap C_c(Y)\).

We say that \((\mathcal{E}, \mathcal{D}(\mathcal{E}))\) admits carré du champ operator if \(\Gamma(u, v) \ll \mathfrak{n}\) for every \(u, v \in \mathcal{D}(\Gamma)\), in which case, with usual abuse of notation, we indicate again by \((\Gamma, \mathcal{D}(\Gamma))\) the \(L^2_n(Y)\)-valued non-negative definite symmetric bilinear form \(\frac{1}{2} \Gamma(u, v)\). By \(\Gamma(u) \leq \mathfrak{n}\) we mean that \(\Gamma(u)\) is absolutely continuous with respect to \(\mathfrak{n}\) and \(\Gamma(u) \leq 1\) \(\mathfrak{n}\)-a.e..

A strongly local Dirichlet form \((\mathcal{E}, \mathcal{D}(\mathcal{E}))\) on \(L^2_n(Y)\) with energy measure \(\Gamma\) induces an intrinsic extended pseudo-metric\(^5\) on \(Y\), termed the intrinsic metric of \((\mathcal{E}, \mathcal{D}(\mathcal{E}))\) and defined by

\[(2.3) \quad d_{\mathcal{E}}(y_1, y_2) := \sup \{u(y_1) - u(y_2) \mid u \in \mathcal{D}(\Gamma) \cap C(Y), \Gamma(u) \leq \mathfrak{n}\} .\]

We will make wide use of the following lemma, which is thus worth to state separately. A proof is standard (see e.g. [28, 2.12] for the first part).

**Lemma 2.3.** Let \((\mathcal{E}, \mathcal{D}(\mathcal{E}))\) be a Dirichlet form on \(L^2_n(Y)\) with energy measure \((\Gamma, \mathcal{D}(\Gamma))\) and let \((u_n)_n \subset \mathcal{D}(\mathcal{E})\) be such that \(\sup_n \mathcal{E}(u_n) < \infty\). If there exists \(u \in L^2_n(Y)\) such that \(L^2_n\)-lim\(n\) \(u_n = u\), then

\[u \in \mathcal{D}(\mathcal{E}) \quad \text{and} \quad \mathcal{E}(u) \leq \liminf_n \mathcal{E}(u_n) .\]

If additionally \((u_n)_n \subset \mathcal{D}(\Gamma)\) and \(\limsup_n \Gamma(u_n) \leq \mathfrak{n}\), then, additionally, \(u \in \mathcal{D}(\Gamma)\) and \(\Gamma(u) \leq \mathfrak{n}\).

**Lemma 2.4.** Assume that \((Y, \tau)\) is additionally compact and let \((\mathcal{E}, \mathcal{D}(\mathcal{E}))\) be a (possibly not regular) strongly local Dirichlet form on \(L^2_n(Y)\) with energy measure \((\Gamma, \mathcal{D}(\Gamma))\). Let \(\rho\) be a metric on \(Y\) metrising the original topology \(\tau\) and assume further that \(\rho_z := \rho(z, \cdot)\) belongs to \(\mathcal{D}(\Gamma)\) and \(\Gamma(\rho_z) \leq \mathfrak{n}\) for every \(z \in Z\) a dense subset of \(Y\).

Then, every \(\rho\)-Lipschitz function \(u: Y \to \mathbb{R}\) satisfies \(u \in \mathcal{D}(\Gamma)\) and \(\Gamma(u) \leq \text{Lip}[u]^2 \mathfrak{n}\).

\(^4\)In the sense of [19, §1.1]. We notice however that, everywhere in the following, we will be interested in Dirichlet forms associated to finite Radon measures on compact Polish spaces, where all common definitions of locality coincide.

\(^5\)By extended we mean that it may attain the value \(+\infty\), by the prefix "pseudo-" that it may vanish outside the diagonal in \(Y \times Y\).

---

RAEMACHER THEOREM ON WASSERSTEIN SPACES

9
PROOF. Without loss of generality, up to rescaling, we can restrict ourselves to the case when \( \text{Lip}[u] \leq 1 \), for which we claim \( \Gamma[u] \leq n \). Let \( u_\varepsilon \) be defined as in Lemma 2.1. Since \( Y \) is compact, functions locally in the domain of the form belong to \( \mathcal{D}(E) \), thus we have \( u_\varepsilon \in \mathcal{D}(E) \) and \( \Gamma(u_\varepsilon) \leq n \) by [24, 2.1] (where the regularity of \( (E, \mathcal{D}(E)) \) is in fact not needed and the fact that \( \Gamma(\rho_n) \leq n \) is granted by assumption). Choose now \( \varepsilon := \varepsilon_n \searrow 0 \) as \( n \to \infty \). Since \( u_{\varepsilon n} \) converges to \( u \) uniformly as \( n \to \infty \) by Lemma 2.1, the conclusion follows by Lemma 2.3. \( \square \)

2.4. Optimal transport. We collect here some known results in metric geometry based on optimal transport. The reader is referred to [2] for an expository treatment.

Everywhere in the following let \( \exp_x : T_x X \to X \) be the exponential map of \( (X, g) \) at a point \( x \in X \) and set \( c := \frac{1}{2} d^2 : X^2 \to \mathbb{R} \).

DEFINITION 2.5 (c-transform, c-convexity, conjugate map). For any \( \varphi : X \to \mathbb{R} \), we define its c-transform\(^6\) by

\[
\varphi^c(x) := - \inf_{y \in M} \{ c(x, y) + \varphi(y) \}.
\]

Any such \( \varphi \) is termed c-convex if there exists \( \psi : X \to \mathbb{R} \) such that \( \varphi = \psi^c \), in which case it holds that \( \varphi = \varphi^{cc} \) (see e.g. [2, 1.9]). Every c-convex function on \( X \) is Lipschitz (see [2, 1.30]\(^7\)). By the classical Rademacher Theorem on \( X \), the set \( \Sigma_{\varphi} \) of singular points of \( \varphi \) has \( m \)-measure 0.

DEFINITION 2.6 (Regular measures). We say that \( \mu \in \mathcal{P} \) is (transport) regular\(^8\) if \( \mu \Sigma_{\varphi} = 0 \) for every semi-convex function \( \varphi \). We denote by \( \mathcal{P}^{\text{reg}} \) the set of regular measures in \( \mathcal{P} \).

REMARK 2.7. The above definition of a regular measure is rather intrinsic. Regularity is a local property. For an extrinsic definition in local charts we refer the reader to [21, 2.8]. The equivalence of our definition to the one in [21] is shown in the proof of [21, 2.10].

THEOREM 2.8 (McCann, [2, 1.33], Gigli, [21, 2.10 and 7.4]). The following are equivalent:

(i) \( \mu \in \mathcal{P}^{\text{reg}} \);

(ii) for each \( \nu \in \mathcal{P} \) there exists a unique optimal transport plan \( \pi \in \text{Opt}(\mu, \nu) \) and \( \pi \) is induced by a map (say, \( g_{\mu \to \nu} \)).

Furthermore, if any of the previous holds, then there exists a c-convex \( \varphi_{\mu \to \nu} \), unique up to additive constant, termed a Kantorovich potential, such that \( g_{\mu \to \nu} = \exp \nabla \varphi_{\mu \to \nu} \) \( \mu \)-a.e. on \( X \).

PROPOSITION 2.9 (AC curves in \( (\mathcal{P}_2, W_2) \), [2, 2.29]). For every \( (\mu_t)_{t \in I} \in AC^1(I; \mathcal{P}_2) \) there exists a Borel measurable time-dependent family of vector fields \( (w_t)_{t \in I} \) such that \( \|w_t\| = |\mu_t| \) for \( dt \)-a.e. \( t \in I \) and the continuity equation

\[
(2.5) \quad \partial_t \mu_t + \text{div}(w_t \mu_t) = 0
\]

\(^6\)Often termed c-transform (e.g. [2]).

\(^7\)The statement, proven in [2] for c-concave functions, is equivalent to our claim by [2, 1.12].

\(^8\)It is well-known that every finite measure on a Polish space is regular in the classical sense of measure theory. Thus we will henceforth refer to ‘transport-regular’ measures simply as to ‘regular’ measures. Since we only consider finite measures on Polish spaces, no confusion may arise.
holds in the sense of distributions on \(I \times X\), that is

\[
\forall \varphi \in C_c^\infty(I \times X) \quad \int_I \int_X d\mu_t(x) \left( \partial_t \varphi(t, x) + \langle \nabla \varphi(t, x) \mid w_t(x) \rangle \right) = 0.
\]

Conversely, if \((\mu_t, w_t)_{t \in I}\) satisfies (2.5) in the sense of distributions and \(\|w_t\|_{\mathcal{X}_\mu} \in L^1(I)\), then, up to redefining \(t \mapsto \mu_t\) on a \(dt\)-negligible set of times, \((\mu_t)_{t \in I} \in AC^1(I; \mathcal{P}_2)\) and \(|\dot{\mu}_t| \leq \|w_t\|_{\mathcal{X}_\mu}\) for \(dt\)-a.e. \(t \in I\).

2.5. Geometry of \(\mathcal{P}_2\). A detailed study of the Riemannian structure of \(\mathcal{P}_2\) has been carried out by N. Gigli in \([21, 22]\), which the present section is mostly inspired by. We shall need the following definitions and results from \([21]\) to which we refer the reader for further references.

We consider the tangent bundle \(TX\) as endowed with the Sasaki metric \(g_s\) and the associated intrinsic distance \(d_s := d_{g_s}\), which turn it into a (non-compact connected oriented) Riemannian manifold.

For \(\mu \in \mathcal{P}_2\) we let \(\mathcal{P}_2(TX)_\mu \subset \mathcal{P}_2(TX)\) be the space of tangent plans \(\gamma \in \mathcal{P}(TX)\) such that

\[
\text{(a)} \quad \text{pr}_T^X \gamma = \mu \quad \text{and} \quad \int_{TX} d\gamma(x, v) |v|^2_{g_s} < \infty.
\]

By \(\exp_\mu : \mathcal{P}_2(TX)_\mu \to \mathcal{P}_2\) we denote the exponential map \(\exp_\mu(\gamma) = \exp_\gamma \gamma\), with right-inverse \(\exp_\mu^{-1} : \mathcal{P}_2 \to \mathcal{P}_2(TX)_\mu\) defined by

\[
\exp_\mu^{-1}(\nu) := \left\{ \gamma \in \mathcal{P}_2(TX)_\mu \mid \exp_\mu(\gamma) = \nu, \int_{TX} d\gamma(x, v) |v|^2_{g_s} = W^2(\mu, \nu) \right\}.
\]

Equivalently, \(\exp_\mu^{-1}(\nu)\) is the set of all tangent plans \(\gamma \in \mathcal{P}_2(TX)\) such that

\[
\text{(b)} \quad (\text{pr}_T^X, \exp)_\gamma \gamma \in \text{Cpl}(\mu, \nu)
\]

and

\[
\text{(c)} \quad \int_{TX} d\gamma(x, v) |v|^2_{g_s} = W^2(\mu, \nu).
\]

Notice that condition (c) may not be dropped (cf. [21, p. 131]), even if (b) is strengthened to

\[
\text{(b') \quad (\text{pr}_T^X, \exp)_\gamma \gamma \in \text{Opt}(\mu, \nu)}.
\]

The joint requirement of both (b) and (c) is however equivalent to that of both (b') and (c).

**Remark 2.10.** Notably, \(\exp_\mu^{-1}(\nu)\) need not be a singleton even when \(\text{Opt}(\mu, \nu)\) is. Consider e.g. the case when \(\mu = \delta_p\) and \(\nu = \delta_q\) are Dirac masses at antipodal points \(p, q \in S^1\) and let \(\nu := \frac{1}{2} \partial_x \in T_p S^1\). Then \(\exp_\mu^{-1}(\nu) = \{\delta_{p, x} + r \delta_{p, -x}\}_{r \in I}\) (cf. (2.1)).

**Proposition 2.11.** Let either \(\mu \in \mathcal{P}_2^{\text{reg}}\) or \(\nu \in \mathcal{P}_2^{\text{reg}}\). Then, \(\exp_\mu^{-1}(\nu)\) is a singleton.

**Proof.** Assume first \(\mu \in \mathcal{P}_2^{\text{reg}}\). By Theorem 2.8 there exists a \(c\)-convex \(\varphi\) (unique up to additive constant) such that

\[
\nu = (\exp. \nabla \varphi.)_\mu \quad \text{and} \quad W^2(\mu, \nu) = \int_X d\mu(x) d^2(x, \exp_x \nabla \varphi_x).
\]

Moreover, for \(\mu\text{-a.e. } x \in X\) there exists a unique geodesic \((\alpha^x_r)_{r \in I}\) connecting \(x\) to \(g_{\mu \rightarrow \nu}(x)\) given by \(\alpha^x_r := \exp_x (r \nabla \varphi_x)\) (cf. [2, 1.35]). We call this property the *geodesic uniqueness* property.
Claim: $\exp_\mu^{-1}(\nu) \neq \emptyset$. Proof. Set $\gamma_0 := (\id_X(\cdot), \nabla \psi)_2 \mu \in \mathcal{P}(TX)$. It is straightforward that $\gamma_0 \in \mathcal{P}_2(TX)_\mu$. Additionally,

\begin{align}
\int_{TX} d\gamma_0(x,v) |\nabla|^2_{g_\nu} &= \int_X d\mu(x) |\nabla \varphi|^2_{g_\nu} \\
&= \int_X d\mu(x) d(x, \exp_x \nabla \varphi)^2 = W^2_2(\mu, \nu),
\end{align}

where $|\nabla \varphi|^2_{g_\nu} = d(x, \exp_x \nabla \varphi)$ for $\mu$-a.e. $x$ by geodesic uniqueness. This shows (c), hence that $\gamma_0 \in \exp_\mu^{-1}(\nu)$.

Claim: $\exp_\mu^{-1}(\nu) = \{\gamma_0\}$. Proof. Let $\gamma \in \exp_\mu^{-1}(\nu)$. By (a), $\pr^X_\mu \gamma = \mu$, thus there exists the Rokhlin disintegration $\{\gamma^x\}_{x \in X}$ of $\gamma$ along $\pr^X_\mu$ with respect to $\mu$. By (b'), $\exp_\mu \gamma = \nu = (\exp. \nabla \varphi)_2 \mu$, thus, for $\mu$-a.e. $x \in X$, $\gamma^x$ is concentrated on the set $A_x := \exp^{-1}_x(\exp_x \nabla \varphi)$. Moreover, (2.8) holds with $\gamma$ in place of $\gamma_0$ by (c), hence, by optimality, $\gamma^x$ is in fact concentrated on the set

\begin{equation}
\pr^{A_x}(0_{T_x X}) := \arg\min_{v \in T_x X} \dist_{g_\nu}(A_x, 0_{T_x X}).
\end{equation}

By geodesic uniqueness, one has $\pr^{A_x}(0_{T_x X}) = \{\nabla \varphi\}$ for $\mu$-a.e. $x \in X$, hence $\gamma^x = \delta_{(x, \nabla \varphi)}$ for $\mu$-a.e. $x \in X$. Thus finally $\gamma = \gamma_0$.

Assume now $\nu \in \mathcal{P}_{\text{reg}}^\nu$. By Theorem 2.8 there exists a $c$-convex $\psi$ (unique up to additive constant) such that (2.7) holds when exchanging $\nu$ with $\mu$ and replacing $\varphi$ with $\psi$. Moreover, geodesic uniqueness holds too, for the geodesics defined by $\beta^\mu_0 := \exp_y(r \nabla \psi_y)$.

For a measurable vector field $w$, we denote by $T^t((\alpha))w_{\alpha_t}$ the parallel transport (of the Levi-Civita connection) from $\alpha_s$ to $\alpha_t$ of the vector $w_{\alpha_t}$ along the curve $((\alpha)) := (\alpha_r)_r$. Set further

\begin{align*}
R &: TX \longrightarrow TX \\
(x, v) &\longmapsto (\exp_x v, -T^1_0((\exp rv)_r)v)
\end{align*}

Claim: $\exp_\mu^{-1}(\nu) \neq \emptyset$. Proof. Set $\gamma_0 := R_4(\id_X(\cdot), \nabla \psi)_2 \mu$. Since

$$\pr^X \circ R \circ (\id_X(\cdot), \nabla \psi) = \exp. \nabla \psi,$$

and $\mu = (\exp. \nabla \psi)_2 \nu$, then $\gamma_0 \in \mathcal{P}_2(TX)_\mu$. Additionally,

\begin{align*}
\int_{TX} d\gamma_0(x,v) |\nabla|^2_{g_\nu} &= \int_X d\nu(y) \left| -T^1_0((\beta^\mu_0)_r) \nabla \psi_y \right|^2_{g_{\beta^\mu_0}} |x := \beta^\mu_0| \\
&= \int_X d\nu(y) |\nabla \psi_y|^2_{g_{\beta^\mu_0}},
\end{align*}

where the last equality holds since, being $(\beta^\mu_0)_r$ a geodesic and the Levi-Civita connection being a metric connection, the parallel transport

$$T^1_0((\beta^\mu_0)_r): (T_{\beta^\mu_0}X, g_{\beta^\mu_0}) \to (T_{\beta^\mu_0}X, g_{\beta^\mu_0})$$

is an isometry. Thus, arguing as in the proof of the first claim, $\gamma_0 \in \exp_\mu^{-1}(\nu)$. 

Claim: $\exp^{-1}_\mu(\nu) = \{\gamma_0\}$. Proof. Let $\gamma \in \exp^{-1}_\mu(\nu)$. By definition $\exp_\mu \gamma = \exp_\nu \gamma = \nu$, thus there exists the Rokhlin disintegration $\{\gamma^y\}_{y \in X}$ of $\gamma$ along $\exp$ with respect to $\nu$. By (b'),

$$(\text{id}_X(\cdot), \exp(-))_\gamma \in \text{Opt}(\mu, \nu) = (\text{pr}^2, \text{pr}^1)_\gamma \text{Opt}(\nu, \mu) = \{(\exp, \nabla \psi, \text{id}_X(\cdot))_\gamma \nu\},$$

thus, for $\nu$-a.e. $y \in X$, $\gamma^y$ is concentrated on the set

$$C_y := \exp_{\beta^y_\nu}^{-1}(y) \subset T_{\beta^y_\nu}X.$$

By a similar reasoning to that in the second claim, for $\nu$-a.e. $y \in X$, $\gamma^y$ is in fact concentrated on $\text{pr}^C(y_{\beta^y_{\nu}} X)$, defined analogously to (2.9). By definition of parallel transport and since $(\beta^y_\nu)_r$ is a geodesic, the latter set is a singleton

$$\text{pr}^C(y_{\beta^y_{\nu}} X) = \{-T^1_0((\beta^y_\nu)_r) \nabla \psi_y\}.$$

This concludes the proof analogously to that of the second claim.

Set further $T^2X := \{(x, v_1, v_2) \mid v_1, v_2 \in T_xX\}$, with natural projections

$$\text{pr}^X: (x, v_1, v_2) \mapsto x \in X, \quad \text{pr}^i: (x, v_1, v_2) \mapsto v_i \in T_xX, \quad i = 1, 2$$

and endowed with the distance

$$d_{t^2} := \sqrt{d^2 \circ (\text{pr}^1, \text{pr}^1) + d^2 \circ (\text{pr}^2, \text{pr}^2)}.$$

For $t \in \mathbb{R}$ we denote by $t \cdot \gamma$ the rescaling

$$t \cdot \gamma := (\text{pr}^X, t\text{pr}^1)_\gamma.$$

Theorem 2.12 (Directional derivatives of the squared Wasserstein distance, [21, 4.2]). Fix $\mu_0 \in \mathcal{P}$ and $\gamma \in \mathcal{P}_2(TX)_{\mu_0}$ and set $\mu_t := \exp_{\mu_0}(t \cdot \gamma)$. Then, for every $\nu \in \mathcal{P}$ there exists the right derivative

$$d^+_\alpha \big|_{t=0} \frac{1}{2} W^2_2(\mu_t, \nu) = -\sup_\alpha \int_{T^2X} d\alpha(x, v_1, v_2) \langle v_1, v_2 \rangle_{g_x}$$

where the supremum is taken over all $\alpha \in \mathcal{P}_2(T^2X)$ such that

$$(\text{pr}^X, \text{pr}^1)_\gamma \alpha = \gamma, \quad \text{and} \quad (\text{pr}^X, \text{pr}^2)_\gamma \alpha \in \exp^{-1}_{\mu_0}(\nu).$$

3. Proof of the main result.

3.1. On the differentiability of $W_2$-cone functions. In this section we collect some results on the differentiability of the Wasserstein distance along (flow) curves. We exploit the fact that, informally, if two flow curves are tangent to each other at every point in the base space $X$, then the lifted (by push-forward) curves on $\mathcal{P}_2$ are themselves, in a sense, tangent to each other.

We denote by $\text{inj}_X > 0$ the injectivity radius of $X$. 
LEMMA 3.1. Let \( w \in X^\infty \). Then,
\[
d(\exp_x(tw_x), \psi^{w,t}(x)) \in o(t) \quad \text{as} \quad t \to 0
\]
uniformly in \( x \in X \).

PROOF. Let \( (\partial_i)_{i=1,\ldots,d} \) be a \( g \)-orthonormal basis of \( T_x X \), \( (d_i)_{i=1,\ldots,d} \) be its \( g \)-dual basis in \( T^*_x X \) and recall the Lie series expansion of \( \psi^{w,t} \) about \( t = 0 \), viz.
\[
\forall f \in C^\infty(X) \quad f(\psi^{w,t}(x)) = \sum_{k \geq 0} \frac{t^k}{k!} w^k(f)_x.
\]

Set \( c_0 := \text{inj}_X(1 \wedge \|w\|_{X^1}^{-1}) \) and let \( 0 < c_1 < c_0 \) be such that \( \psi^{w,t}(x) \in B_{c_1}(x) \) for all \( t < c_1 \).
 Letting \( w_x = w^j \partial_j \) and choosing \( f = d_i \circ \exp_x^{-1} \) (suitably restricted to a coordinate chart around \( x \)) above yields
\[
(d_i \circ \exp_x^{-1})(\psi^{w,t}(x)) = (d_i \circ \exp_x^{-1})(x) + tw(d_i \circ \exp_x^{-1})_x + o(t)
= tw^j \partial_j (d_i \circ \exp_x^{-1}) + o(t) = tw^j + o(t),
\]
whence \( (\exp_x^{-1} \circ \psi^{w,t})(x) = tw + o(t) \). Since \( \exp_x \) is a smooth diffeomorphism on \( B_{c_1}(0_{T_x X}) \), there exists \( L > 0 \) such that
\[
\forall y_1, y_2 \in B_{c_0}(x) \quad d(y_1, y_2) \leq L |\exp_x^{-1}(y_1) - \exp_x^{-1}(y_2)|.
\]
Thus, finally
\[
d(\exp_x(tw), \psi^{w,t}(x)) \leq L |tw - tw - o(t)|_g = o(t),
\]
which concludes the proof. \( \square \)

COROLLARY 3.2. In the same notation of Theorem 2.12, there exists the left derivative
\[
d^-_t \big|_{t=0} \frac{1}{2} W^2_2(\mu_t, \nu) = - \inf_{\alpha} \int_{T^2 X} d\alpha(x, v_1, v_2) \langle v_1 \mid v_2 \rangle_g,
\]
where the infimum is taken over all \( \alpha \in \mathcal{P}_2(T^2 X) \) satisfying (2.12).

PROOF. Given \( \gamma^+ \in \mathcal{P}_2(TX)_{\mu_0} \) let \( \gamma^- := (-1) \cdot \gamma \) be defined by (2.10) and set \( \mu_t^\pm := \exp(\mu \cdot \gamma^\mp) \) for \( t \geq 0 \). Notice that \( \mu_t^- = \mu_t^\pm \) for every \( t \geq 0 \), hence, by definition,
\[
d^-_t \big|_{t=0} \frac{1}{2} W^2_2(\mu_t^+, \nu) = - d^+_t \big|_{t=0} \frac{1}{2} W^2_2(\mu_t^-, \nu)
\]
which exists by choosing \( \gamma = \gamma^- \) in Theorem 2.12. Let \( A^\pm \) be the set of plans \( \alpha \in \mathcal{P}_2(T^2 X) \) satisfying (2.12) with \( \gamma^\pm \) in lieu of \( \gamma \) and define \( r^1 := (pr^X, -pr^1, pr^2); T^2 X \to T^2 X \). It is straightforward that \( A^\pm = r^1 A^\mp \), thus, by Theorem 2.12,
\[
-d^+_t \big|_{t=0} \frac{1}{2} W^2_2(\mu_t^-, \nu) = \sup_{\alpha \in A^-} \int_{T^2 X} d\alpha(x, v_1, v_2) \langle v_1 \mid v_2 \rangle_g
= \sup_{\alpha \in A^+} \int_{T^2 X} d\alpha(x, v_1, v_2) \langle -v_1 \mid v_2 \rangle_g
= - \inf_{\alpha \in A^+} \int_{T^2 X} d\alpha(x, v_1, v_2) \langle v_1 \mid v_2 \rangle_g,
\]
whence the conclusion by combining the last two chains of equalities. \( \square \)
LEMMA 3.3 (Derivatives of the Wasserstein distance along flow curves). Fix \( w \in \mathcal{X}^\infty \), \( \mu_0 \in \mathcal{P} \) and set \( \mu_t := \Psi^{w,t}_\mu \). Then, for every \( \nu \in \mathcal{P} \setminus \{\mu_0\} \), there exists the right derivative

\[
(3.1) \quad d_t^+ |_{t=0} W_2(\mu_t, \nu) = -W_2^{-1}(\mu_0, \nu) \sup_{\gamma} \int_{TX} d\gamma(x,v) \langle w_x | v \rangle_{g_x}
\]

where the supremum is taken over all \( \gamma \in \exp^{-1}_\mu(\nu) \). Moreover, if additionally either \( \mu_0 \in \mathcal{P}^{\text{reg}} \) or \( \nu \in \mathcal{P}^{\text{reg}} \), then there exists the two-sided derivative \( d_t|_{t=0} W_2(\mu_t, \nu) \).

**Proof.** The proof is divided into several steps. Firstly, we show that there exists

\[
(3.2) \quad \lim_{t \downarrow 0} \frac{W_2(\mu'_t, \nu) - W_2(\mu_0, \nu)}{t} = -W_2^{-1}(\mu_0, \nu) \sup_{\gamma} \int_{TX} d\gamma(x,v) \langle w_x | v \rangle_{g_x}
\]

where \( \mu'_t := (\exp(\cdot w \cdot t)) \mu_0 \) and \( \gamma \) is as above. Next, profiting the fact that at \( t = 0 \) the flow \( \exp(tw) \) is tangent to the flow \( \Psi^{w,t}(\cdot) \) at each point in \( X \), we show that the same holds for the corresponding lifted flows \( \exp(tw) \) and \( \Psi^{w,t}(\cdot) \) at each point in \( \mathcal{P} \), hence that the right derivative (3.1) exists and coincides with (3.2).

**Step 1.** Set \( \nu := (\id X \cdot w, w) : X \to TX \), let \( \gamma_0 := \nu^w \mu_0 \in \mathcal{P}_2(TX) \) and notice that

\[
\exp_{\mu_0}(t \cdot \gamma_0) = (\exp_{\mu_0}(pr^X, t pr^X) \circ \nu^w) \mu_0 = (\exp_{\cdot w} \cdot t) \mu_0 =: \mu'_t .
\]

By Theorem 2.12, there exists the right derivative

\[
d_t^+ |_{t=0} \frac{1}{2} W_2^2(\mu'_t, \nu) = -\sup_{\alpha} \int_{T^2X} d\alpha(x,v_1,v_2) \langle v_1 | v_2 \rangle_{g_x}
\]

where \( \alpha \) is as in (2.12). In particular, for every such \( \alpha \), it holds that \( (pr^X, pr^1) \circ \alpha = \gamma_0 = \nu^w |_{\mathcal{P}} \), that is \( (pr^X, pr^1) \circ \alpha \) is supported on the graph \( \text{Graph}(\nu^w) \subset TX \) of the map \( \nu^w \). As a consequence, \( \alpha \) is concentrated on the set

\[
\{(x,v_1,v_2) | (x,v_1) \in \text{Graph}(\nu^w)\} = \{ (x,w_x,v_2) \in T^2X \} \subset T^2X,
\]

thus, in fact

\[
d_t^+ |_{t=0} \frac{1}{2} W_2^2(\mu'_t, \nu) = -\sup_{\gamma} \int_{TX} d\gamma(x,v) \langle w_x | v \rangle_{g_x}
\]

where the supremum is taken over all \( \gamma \in \exp^{-1}_\mu(\nu) \). The existence of \( d_t^+ |_{t=0} W_2(\mu'_t, \nu) \) and (3.2) follow from the existence of \( d_t^+ |_{t=0} W_2(\mu'_t, \nu) \) by chain rule.

**Step 2.** By Lemma 3.1 there exists a constant \( c_1 > 0 \) such that

\[
\forall t \in (0, c_1) \quad \forall x \in X \quad d^2(\exp_{\mu}(tw), \Psi^{w,t}(x)) \in o(t^2).
\]

Furthermore, since \( (\exp_{\cdot w} \cdot t) |_{\mathcal{P}} \mu_0 \) is a coupling between \( \mu'_t \) and \( \mu_t \), equation (3.3) yields

\[
\forall t \in (0, c_1) \quad W_2^2(\mu'_t, \mu_t) \leq \int_X d\mu_0(x) d^2(\exp_{\mu}(tw), \Psi^{w,t}(x)) \in o(t^2),
\]

thus there exists

\[
d_t |_{t=0} W_2(\mu'_t, \mu_t) = \lim_{t \downarrow 0} \frac{1}{t} |W_2(\mu'_t, \mu_t) - W_2(\mu_0, \mu_0)| = 0 .
\]
Step 3. By triangle inequality

\[ W_2(\mu_t, \nu) - W_2(\mu_0, \nu) \leq W_2(\mu_t, \mu'_t) + W_2(\mu'_t, \nu) - W_2(\mu_0, \nu), \]

while by reverse triangle inequality

\[ W_2(\mu, \nu) - W_2(\mu_0, \nu) \geq |W_2(\nu, \mu'_t) - W_2(\mu'_t, \mu)| - W_2(\mu_0, \nu) \geq W_2(\mu'_t, \nu) - W_2(\mu_0, \nu) - W_2(\mu'_t, \mu_t). \]

As a consequence, setting

\[ d^+_t|_{t=0}W_2(\mu_t, \nu) := \limsup_{t \downarrow 0} \frac{W_2(\mu'_t, \nu) - W_2(\mu_0, \nu)}{t}, \]

\[ d^-_t|_{t=0}W_2(\mu_t, \nu) := \liminf_{t \downarrow 0} \frac{W_2(\mu'_t, \nu) - W_2(\mu_0, \nu)}{t}, \]

one has

\[ -d_t|_{t=0}W_2(\mu_t, \nu) + d^+_t|_{t=0}W_2(\mu'_t, \nu) \leq d^-_t|_{t=0}W_2(\mu_t, \nu) \leq d_t|_{t=0}W_2(\mu_t, \mu'_t) + d^+_t|_{t=0}W_2(\mu'_t, \nu) \]

where the derivatives above exist by the previous steps. Since \( d^-_t|_{t=0}W_2(\mu_t, \nu) = 0 \) by Step 2, the right derivative \( d^+_t|_{t=0}W_2(\mu_t, \nu) \) exists and coincides with (3.2).

The last assertion follows by Step 1 and Corollary 3.2 since \( \exp_{\mu_0}^\psi(\nu) \) is a singleton by Proposition 2.11.

\[ \square \]

Lemma 3.4. Let \((X, \mathfrak{g})\) be additionally satisfying assumption (B). Then, for every \( \nu \in \mathcal{P} \) and every \( \theta > 0 \), the function \( u_{\nu, \theta} : \mu \mapsto W_2(\nu, \mu) \lor \theta \) belongs to \( \mathcal{F}_0 \).

Proof. We construct an approximation of \( u_{\nu, \theta} \) by functions in \( \mathcal{F}_C^\infty \).

Preliminaries. By Kantorovich duality (see e.g. [2, 1.17])

\[ W_2^2(\nu, \mu) = 2 \cdot \sup \{ \nu \psi + \mu \varphi \} \]

where the supremum is taken over all \((\psi, \varphi) \in L^1_\nu(X) \times L^1_\mu(X)\) satisfying \( \psi(x) + \varphi(y) \leq c(x, y) \) for \( \nu \)-a.e. \( x \) and \( \mu \)-a.e. \( y \) in \( X \). An optimal pair \((\psi, \varphi)\) always exists and satisfies \( \psi = \varphi^\psi \nu \)-a.e. where \( \varphi^\psi \) is the c-conjugate (2.4) of \( \varphi \).

Let \( \mathcal{P}^{\infty, \times} \) be the set of measures in \( \mathcal{P}^\infty \) with densities bounded away from 0 and fix a countable set \((\mu_i)_i \subset \mathcal{P}^{\infty, \times} \) and dense in \( \mathcal{P}_2 \).

Construction of the approximation. We start by showing that \( W_2(\nu, \cdot) \lor \theta \in \mathcal{F}_0 \) for fixed \( \nu \in \mathcal{P}^{\infty, \times} \). Let \((\psi_i, \varphi_i)\) be the optimal pair of Kantorovich potentials for the pair \((\nu, \mu_i)\), so that

\[ \frac{1}{2}W_2^2(\nu, \mu_i) = \nu \psi_i + \mu_i \varphi_i, \]

where \( \varphi_i \) and \( \psi_i \) are smooth maps by assumption for all \( i \).

Let further \( t = (t_1, \ldots, t_n) \in \mathbb{R}^n \) and, for small \( \varepsilon > 0 \), let \( F_{n, \varepsilon} : \mathbb{R}^n \to [-\varepsilon, \infty) \) be a smooth regularization of the function \( F_n(t) := 2 \cdot \max_{1 \leq i \leq n} t_i \). Since \( F_n \) is 2-Lipschitz for every \( n \), the functions \( F_{n, \varepsilon} \) may be chosen in such a way that (a) \( \lim_{\varepsilon \downarrow 0} F_{n, \varepsilon} = F_n \) on \( \mathbb{R}^n \); (a') \( F_{n, \varepsilon} \) is monotonically
increasing for decreasing $\varepsilon$; and (a’’) $2 \cdot \mathbb{1}_{B_{n,i}} \leq \partial_t F_{n,\varepsilon} \leq 2 \cdot \mathbb{1}_{(B_{n,i})^c}$ for all $i \leq n$, for all $\varepsilon > 0$, for all $n$, where

$$B_{n,i} := \left\{ t \in \mathbb{R}^n \mid t_i > t_j \text{ for all } 1 \leq j < i \right\} + t_i \geq t_j \text{ for all } i \leq j \leq n \right\}$$

are pairwise disjoint and $B_\varepsilon := \{ t \in \mathbb{R}^n \mid \text{dist}(t, B) < \varepsilon \}$ for any $B \subset \mathbb{R}^n$.

For small $0 < \delta < \theta$, let $\varrho_{\theta, \delta}: \mathbb{R} \to [\theta - \delta, \infty)$ be a smooth regularization of $\varrho_\theta: \theta \to \sqrt{1 + \theta}$ such that (b) $\lim_{\delta \downarrow 0} \varrho_{\theta, \delta} = \varrho_\theta$ on $\mathbb{R}$; (b’) $\varrho_{\theta, \delta}$ is monotonically increasing for decreasing $\delta$; (b’’) $\varrho_{(\theta, \infty)}(2\varrho_\theta) \leq \varrho_{\theta, \delta} \leq \varrho_{(\theta - \delta, \infty)}(2\varrho_\theta)$. Now, by smoothness of all functions involved, the function $u_{\theta, n, \varepsilon, \delta}: \mathcal{P} \to \mathbb{R}$ defined by

$$u_{\theta, n, \varepsilon, \delta} (\mu) := \varrho_{\theta, \delta} (F_{n, \varepsilon} (c_1 + \varphi_1^{**} \mu, \ldots, c_n + \varphi_n^{**} \mu)) \quad \text{where} \quad c_i := \psi_i^{**} \nu.$$

belongs to $\mathcal{FC}^\infty$ and one has

$$\nabla u_{\theta, n, \varepsilon, \delta} (\mu) = \sum_{i=1}^{n} \varrho_{\theta, \delta} (F_{n, \varepsilon} (c_1 + \varphi_1^{**} \mu, \ldots, c_n + \varphi_n^{**} \mu) \times$$

$$(\partial_t F_{n, \varepsilon}) (c_1 + \varphi_1^{**} \mu, \ldots, c_n + \varphi_n^{**} \mu) \nabla \varphi_i.$$

By (a) and (a’), resp. (b) and (b’), and Dini’s Theorem, $\lim_{\delta \downarrow 0} \lim_{\varepsilon \downarrow 0} \varrho_{\theta, \delta} \circ F_{n, \varepsilon} (\mathbf{t}) = (\varrho_\theta \circ F_n) (\mathbf{t})$ locally uniformly in $\mathbf{t} \in \mathbb{R}^n$ and for all $n$ and $\theta > 0$. As a consequence, for all $n$ and uniformly in $\mu \in \mathcal{P}$

$$\lim_{\delta \downarrow 0} \lim_{\varepsilon \downarrow 0} u_{\theta, n, \varepsilon, \delta} (\mu) = u_{\theta, n} (\mu) := \varrho_\theta (F_n (c_1 + \varphi_1^{**} \mu, \ldots, c_n + \varphi_n^{**} \mu)) .$$

Moreover, by (a’’), resp. (b’’), $\lim_{\varepsilon \downarrow 0} \partial_t F_{n, \varepsilon} = 2 \cdot \mathbb{1}_{B_{n,i}}$ pointwise on $\mathbb{R}^n$ for all $i \leq n$, for all $n$, resp. $\lim_{\delta \downarrow 0} \varrho_{\theta, \delta} = \mathbb{1}_{[\theta, \infty)} (2\varrho_\theta)$ pointwise on $\mathbb{R}$ for all $\theta > 0$. Thus, for all $n$ and for all $\mu \in \mathcal{P}$ one has

$$\lim_{\delta \downarrow 0} \lim_{\varepsilon \downarrow 0} \nabla u_{\theta, n, \varepsilon, \delta} (\mu) = \sum_{i=1}^{n} \mathbb{1}_{A_{\theta, n, i}} (\mu) \frac{\mathbb{1}_{A_{\theta, n, i}} (\mu)}{\varrho_\theta (F_n (c_1 + \varphi_1^{**} \mu, \ldots, c_n + \varphi_n^{**} \mu))} \nabla \varphi_i,$$

where

$$A_{\theta, n, i} := \left\{ \mu \in \mathcal{P} \mid \begin{array}{l} c_i + \varphi_i^{**} \mu \geq \theta \\
 c_i + \varphi_i^{**} \mu > c_j + \varphi_j^{**} \mu \text{ for all } 1 \leq j < i \\
 c_i + \varphi_i^{**} \mu \geq c_j + \varphi_j^{**} \mu \text{ for all } i \leq j \leq n \end{array} \right\}$$

is measurable by continuity of $\varphi_i^{**}$ for all $i \leq n$, for all $n$.

Finally, again by McCann Theorem, $|\nabla \varphi_i|_g \leq \text{diam} X$, hence

$$|\nabla u_{\theta, n, \varepsilon, \delta} (\mu) (x)|_g \leq n (\text{diam} X) / \sqrt{\theta}$$

whence, by Dominated Convergence, (3.5) and (3.6),

$$\mathcal{E}^{1/2} \lim_{\varepsilon \downarrow 0} (\mathcal{E}^{1/2} \lim_{\delta \downarrow 0} u_{\theta, n, \varepsilon, \delta}) = u_{\theta, n} \in \mathcal{F}_0,$$

$$D u_{\theta, n} (\mu) (x) = \sum_{i=1}^{n} \mathbb{1}_{A_{\theta, n, i}} (\mu) \frac{\mathbb{1}_{A_{\theta, n, i}} (\mu)}{\varrho_\theta (F_n (c_1 + \varphi_1^{**} \mu, \ldots, c_n + \varphi_n^{**} \mu))} \nabla \varphi_i (x) .$$
Pre-compactness of the approximation. Since $L^2_{\mathbb{P}} \lim_n u_{\theta,n} = u_{\nu,\theta}$ by Dominated Convergence and (3.5), by Lemma 2.3 it suffices to show that

$$
(3.7) \quad \text{for } \mathbb{P}\text{-a.e. } \mu \limsup_n \Gamma(u_{\theta,n})(\mu) \leq C_{\nu,\theta}
$$

for some constant $C_{\nu,\theta}$ to get $u_{\nu,\theta} \in \mathcal{F}_0$ and $\Gamma(u_{\nu,\theta}) \leq C_{\nu,\theta}$ $\mathbb{P}$-a.e.. Indeed,

$$
(3.8) \quad \|D u_{\theta,n}(\mu)\|^2_{X_\mu} = \int_X d\mu(x) |D u_{\theta,n}(\mu)(x)|^2_g
$$

since the sets $A_{\theta,n,i}$ are mutually disjoint. Thus

$$
\|D u_{\theta,n}(\mu)\|^2_{X_\mu} = \sum_i \int_X |\varphi_i(\mu)|^2_g \, d\mu \leq \frac{(\text{diam } X)^2}{\theta} =: C_\theta.
$$

General case. Fix an arbitrary $\nu \in \mathcal{P}$ and let $(\nu_k)_k$ be a sequence in $\mathcal{P}^{\infty \times} \mathcal{P}$ narrowly converging to $\nu$. It is readily seen that $u_{\nu_k,\theta}$ converges to $u_{\nu,\theta}$ in $L^2_{\mathbb{P}}(\mathcal{P})$ and $\|D u_{\nu_k,\theta}\|^2_{X_\mu} \leq C_\theta$ $\mathbb{P}$-a.e. by the previous step. Thus, $u_{\nu,\theta} \in \mathcal{F}_0$ and $\|D u_{\nu,\theta}\|^2_{X_\mu} \leq C_\theta$ $\mathbb{P}$-a.e. by Lemma 2.3.

3.2. On the differentiability of functions along flow curves.

Lemma 3.5. Fix $w \in \mathcal{X}^\infty$, $\mu_0 \in \mathcal{P}$ and set $\mu_t := \Psi^{w,t} \mu_0$. Then, the curve $(\mu_t)_{t \in \mathbb{R}}$ is Lipschitz continuous with Lipschitz constant $M \leq \|w\|_{\mathcal{X}_0}$ and satisfies $|\dot{\mu}_t| = \|w\|_{X_{\mu_t}}$ for every $t \in \mathbb{R}$.

Proof. Since constant functions are in particular Lipschitz, we can assume without loss of generality $w \neq 0$. Set $c_1 := \text{inj}_X / \|w\|_{\mathcal{X}_0}$ and let $\mu_{t,\varepsilon} := (\exp_{\varepsilon}(\varepsilon w))_t \mu_t$. For $\varepsilon \in (-c_1, c_1)$, the curve $\varepsilon \mapsto \exp_{\varepsilon}(\varepsilon w)$ is a minimizing geodesic. Thus, $(\exp_{\varepsilon}(\varepsilon w))_t \mu_t \in \text{Opt}(\mu_t, \mu_{t,\varepsilon})$ and, for every $t \in \mathbb{R}$,

$$
d_{\varepsilon} |_{\varepsilon=0} W_2(\mu_t, \mu_{t,\varepsilon}) = \lim_{\varepsilon \to 0} \left( \frac{1}{\varepsilon^2} \int_X d^2(x, \exp_{\varepsilon}(\varepsilon w)) \, d\mu_t(x) \right)^{1/2} = \|w\|_{X_{\mu_t}}.
$$

Arguing as in Step 3 in the proof of Lemma 3.3 with $\mu_{t,\varepsilon}, \mu_{t,\varepsilon}^\prime$ and $\mu_t$ in lieu of $\mu_{t,\varepsilon}, \mu_t$ and $\nu$ respectively,

$$
|\dot{\mu}_t| := d_{\varepsilon} |_{\varepsilon=0} W_2(\mu_t, \mu_{t,\varepsilon}) = d_{\varepsilon} |_{\varepsilon=0} W_2(\mu_t, \mu_{t,\varepsilon}^\prime).
$$

Combining the last two equalities yields the second assertion. Moreover, by [3, 1.1.2],

$$
\forall s \leq t \quad W_2(\mu_s, \mu_t) \leq \int_s^t dr |\dot{\mu}_r| = \int_s^t dr \|w\|_{X_{\mu_r}} \leq \|w\|_{\mathcal{X}_0} |t - s|.
$$

This concludes the proof.

Lemma 3.6. Fix $w \in \mathcal{X}^\infty$, $\mu_0 \in \mathcal{P}$ and set $\mu_t := \Psi^{w,t} \mu_0$. If $u$ is $L$-Lipschitz continuous, then the map $U : t \mapsto u(\mu_t)$ is Lipschitz continuous with $\text{Lip}(U) \leq L \|w\|_{\mathcal{X}_0}$ for every choice of $\mu_0$ and

$$
(3.9) \quad \forall t \in \mathbb{R} \quad |DU(t)| \leq |Du(\mu_t)| \|w\|_{X_{\mu_t}}.
$$
PROOF. The Lipschitz continuity of $U$ follows from those of $u$ and $t \mapsto \mu_t$ (Lem. 3.5). By definition of slope,

$$|DU|(t) \leq \lim_{\nu \to \mu_t} \frac{|u(\mu_t) - u(\nu)|}{W_2(\mu_t, \nu)} \limsup_{s \to t} \frac{W_2(\mu_t, \mu_s)}{|t-s|} = |Du|(\mu_t)|\hat{\mu}_t|$$

for every $t \in \mathbb{R}$, whence (3.9) again by Lemma 3.5.

**Lemma 3.7.** Let $(\mu_t)_{t \in I}$ be an absolutely continuous curve in $\mathcal{P}_2$ connecting $\mu_0$ to $\mu_1$. Then, for every $u \in \mathcal{F}C^\infty$ there exists for a.e. $t \in \mathbb{R}$ the derivative

$$d_tu(\mu_t) = \langle \nabla u(\mu_t) | w_t \rangle_{X_{\mu_t}},$$

where $(\mu_t, w_t)$ is any distributional solution of the continuity equation (2.5), and one has

$$u(\mu_t) - u(\mu_0) = \int_0^t ds \langle \nabla u(\mu_s) | w_s \rangle_{X_{\mu_s}}. \tag{3.10}$$

**Proof.** Let $f$ be in $C^\infty(X)$, $\varphi \in C_c^\infty(\mathbb{R})$ be an arbitrary test function and denote by $\langle \cdot | \cdot \rangle$ the canonical duality pair of distributions. Then,

$$\langle d_tf^{**} \mu_t | \varphi \rangle = \int_\mathbb{R} dt f'(t) f^{**} \mu_t = \int_\mathbb{R} dt \int_X d\mu_t(x) f(x) \varphi'(t)$$

$$= \int_\mathbb{R} dt \int_X d\mu_t(x) \partial_t f(\varphi)(t, x)$$

$$= \int_\mathbb{R} dt \varphi(t) \int_X d\mu_t(x) \langle \nabla f(x) | w_t(x) \rangle_g,$$

for any time dependent vector field $(w_t)_t$ such that $(\mu_t, w_t)_t$ is a solution of (2.5). Thus the distributional derivative is representable by

$$d_t f^{**} \mu_t = \int_X d\mu_t(x) \langle \nabla f(x) | w_t(x) \rangle_g$$

and

$$|d_t f^{**} \mu_t| \leq \|\nabla f\|_{C^0} \|w_t\|_{X_{\mu_t}}.$$

By Proposition 2.9 and absolute continuity of $(\mu_t)_t$ the function $t \mapsto \|w_t\|_{X_{\mu_t}}$ is in $L^1_{loc}(\mathbb{R})$. Thus $t \mapsto d_t f^{**} \mu_t$ is itself in $L^1_{loc}(\mathbb{R})$. Let now $u := F \circ f^{**} \in \mathcal{F}C^\infty$. The above reasoning yields, in the sense of distributions,

$$d_t u(\mu_t) = \sum_{i=1}^k (\partial_i F)(f^{**} \mu_t) d_t f^{**} \mu_t = \sum_{i=1}^k (\partial_i F)(f^{**} \mu_t) \int_X d\mu_t \langle \nabla f_i | w_t \rangle_g,$$

where $(\mu_t, w_t)_t$ is a solution of (2.5) as above and we used (1.3). Since $t \mapsto \nabla u(\mu_t)$ is continuous and bounded by definition of $u$, the distributional derivative of the function $t \mapsto u(\mu_t)$ is again representable by some function in $L^1_{loc}(\mathbb{R})$. Thus, the Fundamental Theorem of Calculus applies and one has

$$u(\mu_t) - u(\mu_0) = \int_0^t ds \langle \nabla u(\mu_s) | w_s \rangle_{X_{\mu_s}} = \int_0^t ds \langle \nabla u(\mu_r) | w_r \rangle_{X_{\mu_r}}.$$

This concludes the proof. \qed
The following two Lemmas are taken — almost verbatim — from [35].

**Lemma 3.8 ([35, 6.1]).** Fix \( w \in \mathcal{X}^\infty \). Then, for every bounded measurable \( u: \mathcal{P} \to \mathbb{R} \) and every \( v \in \mathcal{FC}^\infty \), for every \( t \in \mathbb{R} \)

\[
(3.11) \quad \int_{\mathcal{P}} d\mathbb{P} \left( u \circ \Psi_{w,t} - u \right) v = - \int_0^t ds \int_{\mathcal{P}} d\mathbb{P} \left( u \circ \Psi_{w,s}^{\mu} \right) \nabla_w v.
\]

**Lemma 3.9 ([35, 6.2]).** Fix \( w \in \mathcal{X}^\infty \). Then, for every \( u \in \mathcal{P} \)

\[
(3.12) \quad \forall t \in \mathbb{R} \quad \text{for } \mathbb{P}\text{-a.e. } \mu \quad u(\Psi_{w,t}^\mu) - u(\mu) = \int_0^t dr \left\langle D_u(\Psi_{t}^\mu) \right| w \right\rangle_{\mathcal{X}_{\Psi_{t}^\mu}}
\]

and, for all \( v \in \mathcal{FC}^\infty \),

\[
\forall t \in \mathbb{R} \quad \int_{\mathcal{P}} d\mathbb{P}(\mu) \left\langle D_u(\Psi_{w,t}^\mu) \right| w \right\rangle_{\mathcal{X}_{\Psi_{w,t}^\mu}} v(\mu) = \int_{\mathcal{P}} d\mathbb{P}(\mu) u(\Psi_{w,t}^\mu) \nabla_w v(\mu).
\]

### 3.3. On the differentiability of Lipschitz functions.

In the following let \( u \in \text{Lip}\mathcal{P}_2 \), \( w \in \mathcal{X}^\infty \) and set

\[
(3.13) \quad \Omega_{w}^u := \left\{ \mu \in \mathcal{P} \mid \exists G_w u(\mu) := \frac{d}{dt}|_{t=0} (u \circ \Psi_{w,t}^\mu)(\mu) \right\}.
\]

Since the function \( u \circ \Psi_{w,t}^\mu \) is continuous, the existence of \( G_w u \) coincides with that of the limit \( \lim_{t \to 0} \frac{1}{t}(u(\Psi_{t}^\mu) - u(\mu)) \), \( r \in \mathbb{Q} \). As a consequence the set \( \Omega_{w}^u \) is measurable.

The following proposition, adapted from the proof of [35, 1.3], is at the core of the proof of (3.ii) in our main theorem. Essentially, we prove that, if a Lipschitz function \( u \) on \( \mathcal{P} \) has a directional derivative at some point \( \mu \) for sufficiently many (smooth) directions \( w \), then it is differentiable, in the sense that there exists \( D_u(\mu) \) satisfying the statement of the theorem. This is reminiscent of the same result for Lipschitz functions on \( \mathbb{R}^n \); namely, if \( f: \mathbb{R}^n \to \mathbb{R} \) is locally Lipschitz and Gâteaux differentiable at some point \( x \), then \( f \) is Fréchet differentiable at \( x \) (see [31, Prop. 1]).

**Proposition 3.10.** Fix \( u \in \text{Lip}\mathcal{P}_2 \) and for any \( w \in \mathcal{X}^\infty \) let \( \Omega_{w}^u \) be defined as in (3.13). Let further \( \mathcal{X} \subset \mathcal{X}^\infty \) be a countable \( \mathbb{Q} \)-vector space dense in \( \mathcal{X}^0 \) and assume \( \mathbb{P} \Omega_{w}^u = 1 \) for all \( w \in \mathcal{X} \). Then, the assertions (3) (in particular, (3.i) and (3.ii)) in Theorem 1.4 hold for \( u \).

**Proof.** Fix \( w \in \mathcal{X} \). By assumption on \( \mathcal{X} \), there exists

\[
G_w u(\mu) = \lim_{t \to 0} \frac{u(\Psi_{w,t}^\mu) - u(\mu)}{t}
\]

for all \( \mu \) in the set \( \Omega_{w}^u \) of full \( \mathbb{P} \)-measure. Moreover, by (3.9),

\[
\sup_{t \in [-1,1]} \left| \frac{u(\Psi_{w,t}^\mu) - u(\mu)}{t} \right| \leq \sup_{t \in [-1,1]} \frac{\text{Lip}[u]}{t} \int_0^t dr \left\| w \right\|_{\mathcal{X}_{\Psi_{w,t}^\mu}} \leq \text{Lip}[u] \left\| w \right\|_{\mathcal{X}^0},
\]

thus, by Dominated Convergence,

\[
(3.14) \quad G_w u = L_{\mathbb{P}}^2 \lim_{t \to 0} \frac{u \circ \Psi_{w,t}^\mu - u}{t}.
\]
By continuity of $t \mapsto \frac{1}{t} (u \circ \Psi^{w,t} - u)$, combining Lemma 3.8 with (3.14) yields
\[
\forall v \in \mathcal{F} \mathcal{C}^\infty \quad \int_\mathcal{P} d\mathbb{P} G_w u v = \int_\mathcal{P} d\mathbb{P} u \nabla_w v.
\]

Next, notice that the map $w \mapsto \nabla_w v$ is linear for all $v \in \mathcal{F} \mathcal{C}^\infty$ by assumption ($\mathcal{P}_3$). Hence, if $w = s_1 w_1 + \cdots + s_k w_k$ for some $s_i \in \mathbb{R}$ and $w_i \in \mathcal{X}$, then
\[
\int_\mathcal{P} d\mathbb{P} G_w u v = \sum_{i=1}^k s_i \int_\mathcal{P} d\mathbb{P} u \nabla_{w_i} v = \sum_{i=1}^k s_i \int_\mathcal{P} d\mathbb{P} G_{w_i} u v,
\]
thus

\[(3.15)\quad G_w u = \sum_{i=1}^k s_i G_{w_i} u \quad \mathbb{P}\text{-a.e.}.
\]

Since $\mathcal{X}$ is countable, the set $\bar{\Omega}^\mu := \bigcap_{w \in \mathcal{X}} \Omega^\mu_w$ has full $\mathbb{P}$-measure by assumption. Therefore, the set $\Omega^\mu$ of measures $\mu \in \bar{\Omega}^\mu$ such that $w \mapsto G_w u(\mu)$ is a $\mathcal{Q}$-linear functional on $\mathcal{X}$ has itself full $\mathbb{P}$-measure by (3.15).

For fixed $\mu \in \Omega^\mu$ we have $|G_w u(\mu)| \leq |D u|(\mu) \|w\|_{\mathcal{X}_\mu}$ for every $w \in \mathcal{X}$ by Lemma 3.6. Since $\mathcal{X}$ is $\mathcal{X}_0$-dense in $\mathcal{X}^\infty$, it is in particular $\mathcal{X}_\mu$-dense in $\mathcal{X}^\infty$ for every $\mu \in \mathcal{P}$. Hence the map $w \mapsto G_w u(\mu)$ is a $\mathcal{X}_\mu$-continuous linear functional on the dense subset $\mathcal{X}$ and may thus be extended on the whole space $\mathcal{X}^\infty$ (in fact: on $\mathcal{X}_\mu$) to a continuous linear functional, again denoted by $w \mapsto G_w u(\mu)$ and again such that $|G_w u(\mu)| \leq |D u|(\mu) \|w\|_{\mathcal{X}_\mu}$.

Thus, for every $\mu$ in the set of full $\mathbb{P}$-measure $\Omega^\mu$ there exists $D u(\mu) \in T_\mu \mathcal{P}_2$ such that $G_w u(\mu) = \langle D u(\mu) | w \rangle_{\mathcal{X}_\mu}$ and $\|D u(\mu)\|_{\mathcal{X}_\mu} \leq |D u|(\mu)$. This concludes the proof of the first statement in (3.ii), which in turn implies (3.i) since $|D u|(\cdot) \leq \text{Lip}[u]$.

By definition of $\Omega^\mu$ one has $\Omega^\mu \subset \Omega^\mu_w$ for all $w \in \mathcal{X}$, hence (3.ii) is already proven for all $w \in \mathcal{X}$. In order to prove it for $w \in \mathcal{X}^\infty \setminus \mathcal{X}$, fix $\varepsilon > 0$ and let $w' \in \mathcal{X}$ be such that $\|w - w'\|_{\mathcal{X}_0} < \varepsilon$. Since $X$ is compact, a straightforward modification of [35, 5.5] yields
\[
\|u(\Psi^{w,t} - u(\Psi^{w',t})\| \leq \text{Lip}[u] W_2(\Psi^{w,t}, \Psi^{w',t}) \leq t \text{Lip}[u] c_0 \varepsilon^{c_0 t} \varepsilon
\]
for some constant $c_0 := c_0(X, w) < \infty$. As a consequence,
\[
\forall \mu \in \Omega^\mu \quad \left| \frac{u \circ \Psi^{w,t} - u}{t} - \langle D u(\mu) | w \rangle_{\mathcal{X}_\mu} \right| \leq \varepsilon \text{Lip}[u] c_0 \varepsilon^{c_0 t} + \varepsilon \|D u(\mu)\|_{\mathcal{X}_\mu}
\]
and letting $t \to 0$ yields the conclusion of (3.ii) by arbitrariness of $\varepsilon$.

As consequence of (3.ii) and the bound $\|D u(\mu)\|_{\mathcal{X}_\mu} \leq \text{Lip}[u]$, by definition, $u \in \mathcal{F}_{\text{cont}}$. \hfill \Box

**Corollary 3.11.** Assume $\mathbb{P}$ additionally satisfies ($\mathcal{P}_3$) and let $u \in \text{Lip}\mathcal{P}_2$. Then, the assertions (3.i) and (3.ii) in Theorem 1.4 hold for $u$. 
measure, there exists a sequence $P_2$ for every $w$ with the property of this fact and of Lemma 3.3, Proposition 3.10 applies to the map $F$ again by Lemma 2.3.

Let $\nu \in \mathcal{P}$ be additionally satisfying assumption (B). Then, for every $\nu \in \mathcal{P}$ the function $u_\nu: \mu \mapsto W_2(\nu, \mu)$ belongs to $\mathcal{F}_0$ and $\|D u_\nu\|_X \leq 1$ P.-a.e..

Proof. Assume first $\nu \in \mathcal{P}_{\text{reg}}$ and set $S_\theta(\nu) := \{ \mu \in \mathcal{P} \mid u_\nu(\mu) = \theta \}$. Since $\mathbb{P}$ is a probability measure, there exists a sequence $\theta_n \to 0$ as $n \to \infty$ such that $\mathbb{P} S_\theta(\nu) = 0$. As a consequence of this fact and of Lemma 3.3, Proposition 3.10 applies to the map $u_{\nu, \theta_n} : \mu \mapsto W_2(\nu, \mu) \vee \theta_n$ with $\Omega^{\nu, \theta_n} := \Omega \setminus S_\theta(\nu)$, yielding $\|D u_{\nu, \theta_n}\|_X \leq \text{Lip}[u_{\nu, \theta_n}] = 1$ P.-a.e.. On the other hand, $u_{\nu, \theta_n} \in \mathcal{F}_0$ by Lemma 3.4 and it is clear by reverse triangle inequality that $\lim_n u_{\nu, \theta_n} = u_\nu$ uniformly, whence $u_\nu \in \mathcal{F}_0$ by Lemma 2.3.

If $\nu \in \mathcal{P} \setminus \mathcal{P}_{\text{reg}}$, choose $\nu_n \in \mathcal{P}_{\text{reg}}$ narrowly convergent to $\nu$. Again by reverse triangle inequality $\lim_n u_{\nu, \theta_n} = u_\nu$ uniformly and $\|D u_{\nu, \theta_n}\|_X \leq 1$ P.-a.e. as above, hence the conclusion again by Lemma 2.3.

3.4. Proof of Theorem 1.4.

Proof of (1) AND (2). The proof of [35, 1.4(i) and (iv)], together with the auxiliary results [35, 6.3, 6.4], carries over verbatim to our case. This proves the closability of the forms in assertion (1) and assertion (2). Since $\mathcal{F}_0 \subset \mathcal{F}_{\text{cont}} \subset \mathcal{F}$, it suffices to prove the strong locality of $(\mathcal{E}, \mathcal{F})$. That is, by [6, I.5.1.5] it suffices to show that if $u \in \mathcal{F}$, then $g_1 \circ u, g_2 \circ u \in \mathcal{F}$ and $\mathcal{E}(g_1 \circ u, g_2 \circ u) = 0$ for $g_1, g_2 \in C_c^\infty(\mathbb{R})$ such that $g_1(0) = g_2(0) = 0$ and $\text{supp} g_1 \cap \text{supp} g_2 = \emptyset$.

Fix $w \in X^\infty$ and denote by $\psi_{w,t}(t) \in \mathbb{R}$ its flow. Since $u \in \mathcal{F}$ is bounded, the map $U : t \mapsto u \circ \psi_{w,t}$ satisfies $U(t) \in L^2(\mathcal{P})$ for every $t \in \mathbb{R}$, hence, [35, 6.4] yields for $i = 1, 2$

$$d|_{t=0} g_i(U(t)) = u_i(U(0)) d|_{t=0} U(t) = (g_i \circ u)(D u| w)_X$$

where all derivatives are taken in $L^2(\mathcal{P})$. Hence, the map $\mu \mapsto g_i(u(\mu)) D u(\mu)$ is a measurable section of $T_{\text{Derm}} \mathcal{P}$, satisfies (1.9) and is such that

$$(3.16) \quad \mathcal{E}(g_1 \circ u, g_1 \circ u) = \int \text{d}(\mathcal{P})(u(\mu)) \|D u(\mu)\|_{X^\mu}^2 \leq \|D u\|_{C_0}^2 \mathcal{E}(u, u) < \infty .$$

As a consequence, $g_1 \circ u \in \mathcal{F}$ and the locality property follows now by (3.16) and polarization.

Proof of (3). For fixed $\nu \in \mathcal{P}_{\text{reg}}$ define $u_\nu : \mathcal{P} \to \mathcal{R}$ by $u_\nu : \mu \mapsto W_2(\nu, \mu)$. By Lemma 3.3, for every $\mu \in \Omega^w := \mathcal{P} \setminus \{ \nu \}$ and every $w \in X^\infty$ there exists the limit $G_w u_\nu(\mu)$ defined in (3.13).

Since $\mathbb{P}$ is diffuse by assumption (P2), the set $\Omega^w$ has full $\mathbb{P}$-measure, hence Proposition 3.10 applies to $u_\nu$ with $\Omega^w = \Omega^\nu$ and one has $\|D u_\nu(\mu)\|_{X_\mu} \leq \text{Lip}[u_\nu] = 1$.

Since additionally $u_\nu \in \mathcal{F}_{\text{cont}}$ by Proposition 3.10, if $u$ is $W_2$-Lipschitz continuous, then $u \in \mathcal{F}_{\text{cont}}$ and $\|D u\|_X \leq \text{Lip}[u]$ P.-a.e. by strong locality of $(\mathcal{E}, \mathcal{F})$ and Lemma 2.4 applied to the dense set $\mathcal{P}_{\text{reg}}$, which proves (3.i). If $X$ additionally satisfies assumption (B), then we may replace $\mathcal{F}_{\text{cont}}$ in the above reasoning with $\mathcal{F}_0$ thanks to Corollary 3.12.

If $\mathbb{P}$ additionally satisfies assumption (P4), then assertion (3.ii) reduces to Corollary 3.11.
**Intrinsic distances.** Given a family of functions \( \mathcal{A} \subset \mathcal{F} \) set, for all \( \mu, \nu \in \mathcal{P} \),

\[
d_{\mathcal{A}}(\mu, \nu) \coloneqq \sup \{ u(\mu) - u(\nu) \mid u \in \mathcal{A} \cap \mathcal{C}(\mathcal{P}), \Gamma(u) \leq 1 \text{-a.e. on } \mathcal{P} \} .
\]

**Corollary 3.13 (Intrinsic distances).** Suppose that \( \mathbb{P} \) satisfies assumptions \((P)\) and let \( d_{\mathcal{F}_0} \leq d_{\mathcal{F}_{\text{cont}}} \leq d_{\mathcal{F}} \) be the intrinsic distances \((2.3)\) of the Dirichlet forms \((\mathcal{E}, \mathcal{F}_0)\), \((\mathcal{E}, \mathcal{F}_{\text{cont}})\) and \((\mathcal{E}, \mathcal{F})\) respectively. Then,

\[
d_{\mathcal{F}^\infty} \leq W_2 \leq d_{\mathcal{F}_{\text{cont}}}.
\]

If additionally \((B)\) holds, then the above statement holds with \( d_{\mathcal{F}_0} \) in lieu of \( d_{\mathcal{F}_{\text{cont}}} \).

**Proof.** Let \( \mathcal{A} = \mathcal{F}_0, \mathcal{F}_{\text{cont}}, \mathcal{F} \). If \( u_\nu \in \mathcal{A} \) then

\[
d_{\mathcal{A}}(\mu, \nu) \geq u_\nu(\mu) - u_\nu(\nu) = W_2(\mu, \nu),
\]

hence it suffices to keep track of the assumptions under which \( u_\nu \in \mathcal{F}_0, \mathcal{F}_{\text{cont}}, \mathcal{F} \) respectively in order to show \( W_2 \leq d_{\mathcal{A}} \). One has \( u_\nu \in \mathcal{F}_{\text{cont}} \subset \mathcal{F} \) by the proof of Theorem 1.4(3) above, while \( u_\nu \in \mathcal{F}_0 \) under assumption \((B)\) by Corollary 3.12.

Let now \( u \in \mathcal{F}^\infty \) with \( \|Du\|_{X^\infty} \leq 1 \text{-a.e.} \). Since \( Du = \nabla u \) is continuous, if \( \mathbb{P} \) is fully supported (Assumption \((P_1)\)), then \( \|Du(\mu)\|_{X_\mu} \leq 1 \) for all \( \mu \in \mathcal{P} \). In the same notation of Lemma 3.7, it follows from \((3.10)\) that

\[
u(\mu_1) - u(\mu_0) = \int_0^1 ds \left\langle \nabla u(\mu_s) \mid w_s \right\rangle_{X_{\mu_s}} \leq \int_0^1 ds \|w_s\|_{X_{\mu_s}}.
\]

Taking the infimum of the above inequality over all distributional solutions \((\mu_s, w_s)_{s \in I}\) of \((2.5)\) with fixed \( \mu_0, \mu_1 \) yields \( u(\mu_1) - u(\mu_0) \leq W_2(\mu_0, \mu_1) \) by e.g. \([2, 2.30]\).

This settles all the inequalities in the assertion. \(\square\)

**4. Examples.** Everywhere in this section let \( \phi \in \text{Diff}^\infty(X) \) and denote by \( \Phi : \mathcal{M}_b^+ \to \mathcal{M}_b^+ \) the shift by \( \phi \), by \( \phi^* : \mathcal{L}^0(X) \to \mathcal{L}^0(X) \) the pullback by \( \phi \), and by \( J_\phi^m \) the modulus of the Jacobian determinant of \( \phi \) with respect to \( m \).

Denote further by \( N : \mathcal{M}_b^+ \to \mathcal{P} \) the normalization map \( N : \nu \mapsto \nu = \nu/\nu X \). It is straightforward that \( N \) is continuous with respect to the chosen topologies, hence measurable with respect to the chosen \( \sigma \)-algebras. Moreover, it is readily verified that \( N \) and \( \Phi \) commute, i.e.

\[(4.1)\]

\[N \circ \Phi = \Phi \circ N : \mathcal{M}_b^+ \to \mathcal{P}.
\]

**4.1. On assumption \((P)\).** We collect here some comments on assumption \((P)\). First of all, let us show how one can construct examples of measures satisfying \((P)\) starting from a single one.

**Lemma 4.1.** Let \( w \in X^\infty \) and \( u \in FC^\infty \). Then,

\[
\nabla_w(u \circ \Phi) = \nabla_{\phi \circ u}u \circ \Phi.
\]

**Proof.** Let \( f \in C^\infty(X) \). Then

\[
\nabla(f^* \circ \Phi) = \nabla((f \circ \phi)^*) = \nabla(f \circ \phi).
\]
By (1.4), the proof reduces now to the following computation

\[
\langle \nabla (u \circ \Phi)(\mu) \mid w_x \rangle_{X_\mu} = \sum_{i=1}^{k} (\partial_i F)(f^{**}(\Phi \mu)) \int_X d\mu(x) d(f \circ \Phi)_x(w_x)
\]

\[
= \sum_{i=1}^{k} (\partial_i F)(f^{**}(\Phi \mu)) \int_X d\mu(x) df_{\Phi(x)}(d\Phi_x w_x)
\]

\[
= \sum_{i=1}^{k} (\partial_i F)(f^{**}(\Phi \mu)) \int_X d\Phi(x) df_{\Phi(x)}(d\Phi_{\Phi^{-1}(x)} w_{\Phi^{-1}(x)})
\]

\[
= \langle \nabla u(\Phi \mu) \mid \Phi_* w \rangle_{X_{\Phi \mu}}.
\]

**Proposition 4.2.** Let \( P \in \mathcal{P}(\mathcal{S}) \), \( \phi \in \mathcal{D}^\infty(X) \) and \( \varphi \in \mathcal{F} \) be such that \( \varphi > 0 \) \( P \)-a.e.. Set \( P' := \Phi_2^* P \) and \( P^\varphi := \varphi^2 \cdot P \). Then,

(i) if \( P \) satisfies assumption (P_1), then so do \( P' \) and \( P^\varphi \);

(ii) if \( P \) satisfies assumption (P_2), then so do \( P' \) and \( P^\varphi \);

(iii) if \( P \) satisfies assumption (P_3), then so do \( P' \) and \( P^\varphi \);

(iv) if \( P \) satisfies assumption (P_4), then so does \( P^\varphi \). If additionally \( \phi = \psi^{w,t} \) for some \( w \in X^\infty \), \( t \in \mathbb{R} \), then, additionally, \( P^\varphi \) satisfies assumption (P_4) too.

**Proof.** Since \( \phi \) is bijective, so are \( \Phi := \Phi_2 \) and \( \Phi_4 \). This proves (i) and (ii) for \( P' \); they are also straightforward for \( P^\varphi \), since \( \varphi^2 > 0 \) \( P \)-a.e.. In both cases, (iv) is straightforward by (1.6).

In order to show (iii) for \( P' \), we need to show that there exists an operator \( \nabla_w' : \mathcal{F}C^\infty \to L_2^\varphi(\mathcal{P}) \) such that (1.5) holds with \( P' \) in lieu of \( P \) and \( \nabla_w' \) in lieu of \( \nabla_w \). Since \( \phi \) is a diffeomorphism, the notations \( \phi_*^{-1} \) and \( \phi_1^{-1} = \Phi_1^{-1} = \Phi^{-1} \) are unambiguous. Then, by Lemma 4.1,

\[
\int dP' \nabla_{w'} uv = \int dP \nabla_w u \circ \Phi \cdot v \circ \Phi = \int dP' \nabla_{\phi_*^{-1}w}(u \circ \Phi) v \circ \Phi
\]

\[
= \int dP' u \circ \Phi \cdot \nabla_{\phi_*^{-1}w}(v \circ \Phi) = \int dP' u \nabla_{\phi_*^{-1}w}(v \circ \Phi) \circ \Phi^{-1}.
\]

Assertion (iii) follows by putting \( \nabla_{w'} v := \nabla_{\phi_*^{-1}w}(v \circ \Phi) \circ \Phi^{-1} \).

In order to show (iii) for \( P^\varphi \) assume first that \( \varphi \in \mathcal{F}C_\infty \), whence \( \varphi \) is continuous and bounded (Rmk. 1.2). Then, by (1.3) and (1.4)

\[
\int dP u\varphi^2 \nabla_w v = \int dP \nabla_{w}(u\varphi^2) \cdot v = \int dP \nabla_w u \cdot \varphi^2 v + \int dP \varphi^2 uv \cdot (2\varphi^{-1} \nabla_w \varphi)
\]

and the assertion follows by setting \( \nabla_{w}^{\varphi^2} v := \nabla_w v - (2\varphi^{-1} \nabla_w \varphi) v \). The general case follows by approximation as soon as we show that the pre-Dirichlet form

\[
\mathcal{F}^\varphi := \left\{ u \in \mathcal{F} \mid \int_{\mathcal{F}} dP \varphi^2(u^2 + ||Du||^2_{X}) < \infty \right\}
\]

\[
\mathcal{E}^\varphi(u,v) := \int_{\mathcal{F}} dP \varphi^2(Du \mid Dv)_{X}
\]

is closable. Provided that \( (\mathcal{E}, \mathcal{F}) \) is a strongly local Dirichlet form by Theorem 1.4(1), this last assertion is the content of [14, 1.1].
Remark 4.3. While points (i)–(iii) of the Proposition suggest that assumptions (P1)–(P3) are quite generic with respect to shifting \( P \) by (the lift of) a diffeomorphism, point (iv) is (by far) more restrictive, as the inclusion \( \text{Flow}(X) \subseteq \text{Diff}_+^r(X) \) is always strict, even on \( S^1 \), see e.g. [23].

It is clear that the closability of the pre-Dirichlet forms \((\mathcal{E}, \mathcal{F}^\infty)\) and \((\mathcal{E}, \mathcal{F}_{\text{cont}})\) associated to \( P \) is essential to our approach in discussing Rademacher-type theorems, which settles the necessity of assumption (P3). The necessity of assumption (P1) is instead motivated by the following trivial example.

Example 4.4. Denote by \( \delta: X \mapsto \mathcal{P} \) the Dirac embedding \( x \mapsto \delta_x \) and set \( \mathcal{P} := \delta \sharp \mathbb{m} \). Since \( \mathcal{P} \) is supported on the family of Dirac masses, it does not satisfy (P1). On the other hand, since \( W_2(\delta_{x_1}, \delta_{x_2}) = d_g(x_1, x_2) \) for every \( x_1, x_2 \in X \), it is also clear that \((\mathcal{P}, W_2, \mathcal{P})\) and \((X, d_g, \mathbb{m})\) are isomorphic as metric measure spaces, which shows (P2). Moreover,

\[
\Phi \circ \delta \sharp \mathbb{m} = (\Phi \circ \delta)_\sharp \mathbb{m} = (\delta \circ \phi)_\sharp \mathbb{m} = \delta_\sharp (\Phi \mathbb{m}) = \delta_\sharp (J^{\mathbb{m}}_\Phi \cdot \mathbb{m}) = (J^{\mathbb{m}}_\Phi)^{**} \cdot \delta_\sharp \mathbb{m}
\]

and (P3) holds for \( P \) as well.

Remark 4.5. Incidentally, notice that Theorem 1.4 applied to Example 4.4 provides a non-local proof of the classical Rademacher Theorem on a closed Riemannian manifold. Indeed it suffices to notice that \( T^{\text{Der}}_\delta X = T_X X \) as Hilbert spaces for every \( x \in X \) and that every Lipschitz function \( f \in \text{Lip}(X) \) induces a Lipschitz function \( \tilde{f} \in \text{Lip}(\mathcal{P}_2) \), namely the (e.g. lower) McShane extension \( \tilde{f} \) of the function \( f \circ \delta^{-1} \) defined on the image of \( \delta \).

Assumption. For \( \mathcal{P} \) a Borel probability measure on \( \mathcal{P}_2 \) set:

(P5) assumption (P4) holds and the Radon-Nikodým derivative \( R^\mathcal{P}_r \) defined in (1.6) is such that for every \( w \in \mathcal{X}^\infty \):

- \( r \mapsto R^\mathcal{P}_r(\mu) \) is differentiable in a neighborhood of 0 for \( \mathcal{P} \)-a.e. \( \mu \);
- \( \mu \mapsto |\partial_r R^\mathcal{P}_r(\mu)| \) is integrable w.r.t. \( \mathcal{P} \) uniformly in \( r \) on a neighborhood of 0.

Proposition 4.6. The following chain of implications holds true:

\[
(P5) \implies (P4) \land (P3) \implies (P4) \implies (P2).
\]

In particular: (P1) \land (P5) \implies (P).

Proof. The implication \( (P5) \implies (P4) \) is trivial and it is readily seen that assumption (P2) is already implied by the first part of (P4). It remains to show that \( (P5) \implies (P3) \). Indeed,

\[
\int_{\mathcal{P}} \text{d}\mathcal{P} \nabla_w u \cdot v = \int_{\mathcal{P}} \text{d}\mathcal{P} \lim_{t \to 0} \frac{u \circ \Psi^{w,t} - u}{t} \cdot v = \lim_{t \to 0} \frac{1}{t} \int_{\mathcal{P}} \text{d}\mathcal{P} (u \cdot v \circ \Psi^{w,-t} \cdot R^w_{-t} - uv) = \lim_{t \to 0} \frac{1}{t} \int_{\mathcal{P}} \text{d}\mathcal{P} u (v \circ \Psi^{w,-t} - v) + \lim_{t \to 0} \frac{1}{t} \int_{\mathcal{P}} \text{d}\mathcal{P} u (v \circ \Psi^{w,-t}) (R^w_{-t} - 1)
\]
\begin{align*}
+ \lim_{t \to 0} \frac{1}{t} \int_{\mathcal{P}} \mathbb{D}^{\mathbb{P}} u v \left( R_{w-t}^w - 1 \right).
\end{align*}

The first limit in the last equality satisfies, by Dominated Convergence,
\begin{align*}
\lim_{t \to 0} \frac{1}{t} \int_{\mathcal{P}} \mathbb{D}^{\mathbb{P}} u (v \circ \Psi_{w-t}^w - v) = - \int_{\mathcal{P}} \mathbb{D}^{\mathbb{P}} u \cdot \nabla_w v.
\end{align*}

The second limit vanishes, again by Dominated Convergence, since \( t \to R_t^w (\mu) \) is continuous (differentiable) at \( t = 0 \) for \( \mathbb{P} \)-a.e. \( \mu \). In light of assumption \((P3)\), differentiating under integral sign, the third limit satisfies
\begin{align*}
\lim_{t \to 0} \frac{1}{t} \int_{\mathcal{P}} \mathbb{D}^{\mathbb{P}} u v (R_{w-t}^w - 1) = \int_{\mathcal{P}} \mathbb{D}^{\mathbb{P}} u v \cdot \partial_t |_{t=0} R_{w-t}^w.
\end{align*}

As a consequence, assumption \((P3)\) is satisfied by letting
\begin{align*}
\nabla_w^w v := - \nabla_w v - \partial_t |_{t=0} R_{w-t}^w \cdot v.
\end{align*}

This concludes the proof. \( \square \)

4.2. On assumption \((B)\). The reader is referred to [17] and references therein for an expository treatment of regularity theory of optimal transport maps on Riemannian manifolds, whereof we make use in the present section. We denote by \( \text{ST}_2X := \{ w \in T_xX \mid |w|_{g_x} = 1 \} \) the unit tangent space to \( (X, g) \) at \( x \). Everywhere in the following also let \( c := \frac{1}{2} d^2 \).

Further geometrical assumptions. For \( x \in X \) and \( w \in T_xX \) define the cut, resp. focal, time by
\begin{align*}
t_c(x, w) := \inf \{ t > 0 \mid s \mapsto \exp_x(sw) \text{ is not a } d\text{-minimizing curve from } x \text{ to } \exp_x(tw) \}
\end{align*}
\begin{align*}
t_f(x, w) := \inf \{ t > 0 \mid \text{d}_{tw} \exp_x : T_xX \to T_{\exp_x(tw)}X \text{ is not invertible} \}
\end{align*}
and the (tangent), resp. (tangent) focal, cut locus and injectivity domain by
\begin{align*}
TCL(x) := \{ t_c(x, w)w \mid w \in \text{ST}_2X \}, \\
TFL(x) := \{ t_f(x, w)w \mid w \in \text{ST}_2X \}, \\
I(x) := \{ tw \mid w \in \text{ST}_2X, 0 \leq t < t_c(x, w) \}.
\end{align*}

Finally, recall the definition of the Ma–Trudinger–Wang tensor
\begin{align*}
\mathcal{S}(x, y)(w, w') := - \frac{3}{2} \left. d^2 \right|_{s=0} \left. d^2 \right|_{t=0} c(\exp_x(tw), \exp_x(p + sw)),
\end{align*}
where \( x \in X, y \in I(x), w, w' \in T_xX \) and \( p := \exp_x^{-1}(y) \).

The following definitions are taken from [17].

DEFINITION 4.7 (Non-focality of cut loci). We say that \( (X, g) \) is non-focal if it additionally satisfies \( \text{fcut}(x) = \emptyset \) for all \( x \in X \).

DEFINITION 4.8 (Strong Ma–Trudinger–Wang condition \( \text{MTW}(K) \)). We say that \( (X, g) \) satisfies the strong Ma–Trudinger–Wang condition with constant \( K \) (in short: \( X \) is \( \text{MTW}(K) \)) if there exists a constant \( K > 0 \) such that
\begin{align*}
\forall x \in X, y \in \exp_x(I(x)) \quad \mathcal{S}(x, y)(w, w') \geq K |w|_{g_x}^2 |w'|_{g_x}^2 \text{ whenever } w^* \left[ c_{\ldots}, \right] w' = 0,
\end{align*}
where \( [c_{\ldots}] \) denotes the matrix of derivatives \( c_{i,j} := \partial^2_{x_i, y_j} c \).
Our main interest in the previous definitions is the following regularity result.

THEOREM 4.9 (Loeper–Villani (See e.g. [17, 3.13].)). Let \((X, \mathbf{g})\) be additionally non-focal and satisfying MTW\((K)\). Then \(X\) satisfies assumption \((\mathbf{B})\).

REMARK 4.10. The strong MTW condition is sufficient, whereas not necessary, to establish the above result. A discussion of optimal assumptions is here beyond our purposes. It will suffice to say that the proof strategy of Lemma 3.4 fails as soon as \(c\)-convex \(C^1\) functions are not uniformly dense in (Lipschitz) \(c\)-convex functions (see [17, 3.4]).

4.3. Normalized mixed Poisson measures. We denote by \(\tilde{\Gamma}\) the space of integer-valued Radon measures over \((X, \mathbf{g})\) with arbitrary finite number of atoms, always regarded as a subspace of \(\mathcal{M}^+_b\), endowed with the vague topology (which coincides with the narrow topology by compactness) and with the associated Borel \(\sigma\)-algebra. Similarly to [1, 35], we let \(\rho \in C^1(X; \mathbb{R}^+)\) and denote by \(\mathcal{P}_\sigma\) the Poisson measure of intensity \(\sigma := \rho \mathbf{m}\) on \(\tilde{\Gamma}\). Given \(\lambda \in \mathcal{P}(\mathbb{R}^+)\) such that \(\lambda(1 \land \text{id}_{\mathbb{R}^+}) < \infty\), henceforth a Lévy measure, we denote by \(\mathcal{R}_{\lambda, \sigma}\) the mixed Poisson measure \(\mathcal{R}_{\lambda, \sigma} = \int_{\mathbb{R}^+} d\lambda(s) \mathcal{P}_{s, \sigma}\). Recall that \(\mathcal{P}_\sigma\), hence \(\mathcal{R}_{\lambda, \sigma}\), is concentrated on the configuration space

\[
\Gamma := \left\{ \gamma \in \tilde{\Gamma} \mid \gamma \{x\} \in \{0, 1\} \text{ for all } x \in X \right\}.
\]

Moreover (see [1, Prop. 2.2]), for all \(\gamma \in \Gamma\)

\[
(\Phi^t \mathcal{P}_\sigma) (\gamma) = \exp (\sigma(1 - p^\gamma_b)) \prod_{x \in \gamma} p^\gamma_b(x) , \quad \text{where} \quad p^\gamma_b(x) := \frac{\Phi^s \rho(x)}{\rho(x)} J^s_\gamma (x)
\]

and by \(x \in \gamma\) we mean \(\gamma \{x\} > 0\). Since we chose \(\rho \in L^1_m(X)\), the measure \(\sigma\) is finite, hence \(\gamma X < \infty\) for \(\mathcal{P}_\sigma\)-a.e. \(\gamma\), i.e. \(\mathcal{P}_\sigma\)-a.e. \(\gamma\) is concentrated on a finite number of points. As a consequence, the same statement holds for \(\mathcal{R}_{\lambda, \sigma}\) in lieu of \(\mathcal{P}_\sigma\) and one has

\[
R^\gamma (\gamma) := \prod_{x \in \gamma} p^\gamma_b(x) = \exp \int_X \gamma(x) \ln (p^\gamma_b(x)) .
\]

EXAMPLE 4.11 (Normalized mixed Poisson measures). Let \(\lambda \in \mathcal{P}(\mathbb{R}^+)\) be a Lévy measure with compact support and set \(\mathbb{P} := N_1 R_{\lambda, \sigma}\). Assumption \((P_2)\) is satisfied because of the diffuseness of \(\sigma\), whence that of \(\mathcal{P}_\sigma\) and, in turn, that of \(R_{\lambda, \sigma}\). Assumptions \((P_4)\) and \((P_3)\) are respectively verified in Lemmas 5.5 and 5.6 below. In particular, the closability of the pre-Dirichlet form in (1.11) is obtained as a consequence of the quasi-invariance of \(\mathbb{P}\). Assumption \((P_1)\) is verified in Lemma 5.7 below.

Denote now by \(X^{\otimes n} := X^{\times n} / \mathcal{S}_n\) the quotient of the \(n\)-fold cartesian product \(X^{\times n}\) by the symmetric group \(\mathcal{S}_n\) acting by permutation of coördinates. Let further \(\tilde{X}^{\times n}\) denote the set of points \(x := (x_1, \ldots, x_n) \in X^{\times n}\) such that \(x_i \neq x_j\) for \(i \neq j\), and set

\[
X^{(n)} := \tilde{X}^{\times n} / \mathcal{S}_n.
\]

Denote by \(pr^{\otimes n} : \tilde{X}^{\times n} \to X^{(n)}\) the quotient projection, and set \(\sigma^{(n)} := pr^{\otimes n}_\sigma \sigma^{\otimes n}\). It is well-known that, when \((X, \sigma)\) is a finite Radon measure space, then \((\Gamma, \mathcal{P}_\sigma)\) is isomorphic, as a
measure space, to the space

\[ \bigoplus_{n \in \mathbb{N}_1} \left( X^{(n)}, e^{-\sigma X \sigma^{(n)}/n!} \right). \]

More explicitly, the isomorphism is given by identifying \( X^{(n)} \) with \( \Gamma^{(n)} \), the space of configurations \( \gamma \in \Gamma \) such that \( \gamma X = n \). Finally, define the following subsets of \( \mathcal{P} \)

\[
N(\Gamma^{(n)}) \subset \Delta^n := \left\{ \sum_{i} s_i \delta_{x_i} \mid x \in X^{(n)}, s_i \in \mathbb{R}^+ \right\} \subset \tilde{\Delta}^\text{fin} := \bigcup_n \Delta^n,
\]

\[
N(\tilde{\Gamma}^{(n)}) \subset \Delta^n := \left\{ \sum_{i} s_i \delta_{x_i} \mid x \in X^{\times n}, s_i \in \mathbb{R}^+ \right\} \subset \Delta^\text{fin} := \bigcup_n \Delta^n.
\]

\textbf{Remark 4.12.} While the support \( \tilde{\Delta}^1 = \Delta^1 \cong X \) of the measure constructed in Example 4.4 is “small” in various senses — e.g., it is a closed nowhere dense subset of \( \mathcal{P} \), the normalized (mixed) Poisson measures in Example 4.11 are fully supported. On the other hand, though, even these measures are concentrated on \( \tilde{\Delta}^\text{fin} \), which may itself be still regarded as “small” — e.g., since the measure space \( (\tilde{\Delta}^\text{fin}, N_\mathcal{R}_{\sigma,\lambda}) \) may be approximated in many senses via the sequence of compact finite-dimensional measure spaces \( (\Delta^n, N_\mathcal{R}_{\sigma,\lambda}|_{\Delta^n}) \).

\textbf{4.4. The Dirichlet–Ferguson measure.} Example 4.11 shows that the laws of (normalized) point processes on \( X \) may be examples of measures on \( \mathcal{P} \) satisfying assumption (P). In light of Remark 4.12, the question arises, whether such laws may be chosen to be concentrated on sets richer than \( \tilde{\Delta}^\text{fin} \), and in particular on the whole set of purely atomic measures.

In this section we introduce for further purposes a negative example, the \textit{Dirichlet–Ferguson measure}. The unacquainted reader may take this result as a definition of the Dirichlet–Ferguson measure.

\textbf{Preliminaries.} Denote by \( \mathfrak{m} \) the \textit{normalized} volume measure of \( X \). Everywhere in the following let \( \beta \in (0, \infty) \) be defined by \( m = \beta \mathfrak{m} \). Set further \( \hat{X} := X \times I \), always endowed with the product topology, \( \sigma \)-algebra and with the measure \( m_\beta := \mathfrak{m} \otimes B_\beta \), where

\[ dB_\beta(r) := \beta(1-r)^{\beta-1} \, dr \]

is the Beta distribution on \( I \) with parameters \( 1 \) and \( \beta \).

\textbf{The Dirichlet–Ferguson measure.} We denote by \( \mathcal{D}_m \) the Dirichlet–Ferguson measure \cite{16} over \( (X, B) \) with intensity \( m \). The characteristic functional of \( \mathcal{D}_m \) may be found in [12] together with the further properties of the measure. The following characterization is originally found, in the form of a distributional equation, in [36, (3.2)].

\textbf{Theorem 4.13 (Mecke-type identity for \( \mathcal{D}_m \) \cite{36}, see also \cite{13}).} Let \( u \colon \mathcal{P} \times \hat{X} \to \mathbb{R} \) be measurable semi-bounded. Then, there exists a unique measure \( \mathcal{D}_m \) on \( \mathcal{P} \) satisfying

\[
\int_{\mathcal{P}} d\mathcal{D}_m(\eta) \int_X d\eta(x) u(\eta, x, \eta x) = \int_{\mathcal{P}} d\mathcal{D}_m(\eta) \int_X d\mathfrak{m}_\beta(x, r) u(\eta + r \delta_x, x, r).
\]

The unacquainted reader may take this result as a definition of \( \mathcal{D}_m \).

---

\( ^9 \) Among the many other names: \textit{Dirichlet}, Poisson–Dirichlet \cite{40}, (law of the) Fleming–Viot process with parent-independent mutation \cite{33}. 

---
4.5. The entropic measure. In this section we recall an example showing that there exist measures on $\mathcal{P}$ — other than normalized mixed Poisson measures — satisfying assumptions (P).

Preliminaries. Similarly to [34, §2.2], define

$$\mathcal{G}(\mathbb{R}) := \{ g: \mathbb{R} \to \mathbb{R}, \text{ right-cont. non-decr., s.t. } \forall x \in \mathbb{R} \quad g(x + 1) = g(x) + 1 \}$$

In light of the equi-variance property, each $g \in \mathcal{G}(\mathbb{R})$ uniquely induces a Borel function $\text{pr}^\mathcal{G}(g): \mathbb{S}^1 \to \mathbb{S}^1$ and we set $\mathcal{G} := \text{pr}^\mathcal{G}(\mathcal{G}(\mathbb{R}))$, endowed with the $L^2$-distance

$$\|g_1 - g_2\|_{\mathcal{G}} := \left( \int_{\mathbb{S}^1} \text{dm}(t) \ |g_1(t) - g_2(t)|^2 \right)^{1/2}.$$

Letting $\mathbb{S}^1 \cong \mathbb{R} / \mathbb{Z}$, define further for every $a \in \mathbb{S}^1$ the translation $\tau_a: \mathbb{S}^1 \to \mathbb{S}^1$ by

$$\tau_a: t \mapsto t + a \pmod{1},$$

and define an equivalence relation $\sim$ on $\mathcal{G}$ by setting $g \sim h$ for $g, h \in \mathcal{G}$ if and only if $g = h \circ \tau_a$ for some $a \in \mathbb{S}^1$. Denote by $\text{pr}^\mathcal{G}_1$ the quotient map of $\mathcal{G}$ modulo this equivalence relation, with values in the quotient space $\mathcal{G}_1 := \text{pr}^\mathcal{G}_1(\mathcal{G}) = \mathcal{G}/\mathbb{S}^1$ endowed with the quotient $L^2$-distance

$$\|g_1 - g_2\|_{\mathcal{G}_1} := \left( \inf_{s \in \mathbb{S}^1} \int_{\mathbb{S}^1} \text{dm}(t) \ |g_1(s) - g_2(t + s)|^2 \right)^{1/2}.$$

Equivalently, $\mathcal{G}_1$ is the semi-group of right-continuous non-decreasing functions on $\mathbb{S}^1 \cong [0, 1)$ fixing $0 \in \mathbb{S}^1$. Finally, the space $(\mathcal{G}_1, \| \cdot \|_{\mathcal{G}_1})$ is isometric (see [34, 2.2]) to $\mathcal{P}_2 := (\mathcal{P}(\mathbb{S}^1), W_2)$ via the map

$$\chi: g \mapsto g^\sharp \mu.$$

The conjugation map $\mathcal{G}^\sharp$ (cf. [37, §3]). For $\mu \in \mathcal{P}$ let $\varphi_\mu := \varphi_{\mathcal{G} \rightarrow \mu}$ be given by Theorem 2.8 (recall that $\mathcal{G} \in \mathcal{P}_{\text{reg}}$). The conjugation map $\mathcal{G}^\sharp: \mathcal{P} \to \mathcal{P}$ is defined by

$$\mathcal{G}^\sharp: \mu \mapsto (\exp \nabla (\varphi_\mu))_{\sharp} \mathcal{G}.$$

It was shown in [37, 3.5] that $\mathcal{G}^\sharp$ is an involutive homeomorphism of $\mathcal{P}_2$. If $X = \mathbb{S}^1$, then the conjugation map may be alternatively defined in the following equivalent way. Let

$$g_\mu(t) := \inf \{ s \in I \mid \mu[0, s] > t \} \quad (\text{here: conventionally, } \inf \emptyset := 1)$$

denote the cumulative distribution function of $\mu \in \mathcal{P}(\mathbb{S}^1)$. Observe that $g_\mu \in \mathcal{G}_1$, hence it admits a left inverse $g_\mu^{-1}$ in $\mathcal{G}_1$ given by

$$g_\mu^{-1}(t) := \inf \{ s \geq 0 \mid g(s) > t \}.$$

Then, $\mathcal{G}^\sharp(\mu) = dg_\mu^{-1}$ where, for any $g \in \mathcal{G}_1$, we denoted by $dg$ the Lebesgue–Stieltjes measure associated to $\varphi$ (see [34] for the detailed construction).

**Definition 4.14 (entropic measure over $X$ [37, 6.1]).** The entropic measure $\mathcal{P}_m$ is the Borel probability measure on $\mathcal{P}_2$ defined by $\mathcal{P}_m := \mathcal{G}^\sharp \mathcal{D}_m$, where $\mathcal{D}_m$ is the Dirichlet–Ferguson measure of §4.4.
Since $\mathcal{C}^m$ is a homeomorphism, $\mathbb{P}_m$ satisfies assumptions $(P_1)$, $(P_2)$ because so does $D_m$. The quasi-invariance of $\mathbb{P}_m$ as in assumption $(P_4)$ and assumption $(P_3)$ (hence the closability of the Dirichlet form (1.11)) are a challenging problem. They have been proven in the seminal work [34] for the case $X = S^1$, which leads us to the following example.

Example 4.15 (The entropic measure over $S^1$ [34, 3.3]). Let $\beta > 0$ be a fixed constant and let $X = S^1$ be endowed with the rescaled volume measure $m := \beta \mathcal{L}^1$. The quasi-invariance of $\mathbb{P}_m$ as in assumption $(P_4)$ was proven in [34, 4.2] (in fact, it was proven for the action of the whole of $\text{Diff}^2(X)$ rather than only for $\text{Flow}(X)$, cf. Rmk. 4.3). Although not apparent, the bound (1.6) for the Radon–Nikodym derivative $R_{\mathbb{P}}^m$ may be deduced from the explicit computations in [34, 4.8]. In fact, assumption $(P_3)$ holds too, because of [34, 5.1(ii)]. Assumption $(P_3)$ holds as a consequence of $(P_5)$ by Proposition 4.6. Together with the previous discussion, this shows that $\mathbb{P}_m$ satisfies assumption $(P)$.

The closability of the form $(\mathcal{E}, \mathfrak{F}_0)$ is proven in [34, 7.25], a proof of the Rademacher property in the form of our Theorem 1.4(3.i) is sketched in [34, 7.26].

Remark 4.16. Finally, let us notice that $\mathbb{P}_m$-a.e. $\mu$ is concentrated on an $m$-negligible set [34, 3.11]. In fact, it is not difficult to show that $\mathbb{P}_m$-a.e. $\mu$ is concentrated on the set of irrational points of a Cantor space.

4.6. An image on $\mathcal{P}$ of the Malliavin–Shavgulidze measure. As a final example, we introduce here an image on $\mathcal{P}(S^1)$ of the Malliavin–Shavgulidze measure on $\text{Diff}^1_+(S^1)$.

Preliminaries. See e.g. [26] for a detailed exposition and further references. Let $X = S^1$ with volume measure $m := \mathcal{L}^1$ and denote by $C_0(I^0)$ the space of continuous functions on $I$ vanishing at both 0 and 1, endowed with the trace topology of $C(I)$. Consider the space $\text{Diff}^1_+(S^1)$ of orientation preserving $C^1$-diffeomorphisms of $S^1$, endowed with the topology of uniform convergence, and let $\xi: \text{Diff}^1_+(S^1) \to S^1 \times C_0(I^0)$ be the homeomorphism defined by

$$\xi: g(t) \mapsto (g(0), \ln g'(0) - \ln g'(0)).$$

Definition 4.17 (The Malliavin–Shavgulidze measure). Let $\mathcal{W}_0$ be the Borel probability on $C(I)$ defined as the law of the Brownian Bridge connecting 0 to 0 in time 1, concentrated on $C_0(I^0)$. The Malliavin–Shavgulidze measure $\mathcal{M}$ on $\text{Diff}^1_+(S^1)$ [29] is the Borel probability measure defined by $\mathcal{M} := (\xi^{-1})_!(m \otimes \mathcal{W}_0)$.

Denote further by $S$ the Schwarzian derivative operator

$$S: \phi \mapsto \frac{\phi'''}{\phi'} - \frac{3}{2} \left( \frac{\phi''}{\phi'} \right)^2,$$

and consider the left action $L_\phi: g \mapsto \phi \circ g$ of the subgroup $\text{Diff}^1_+(S^1)$. The measure $\mathcal{M}$ is quasi-invariant with respect to $L_\phi$ and the following quasi-invariance formula holds true (see e.g. [29]) for every Borel $A \subset \text{Diff}^1_+(S^1)$

$$\mathcal{M}(L_\phi(A)) = \int_A d\mathcal{M}(g) \exp \left[ \int_{S^1} dm(t) S(\phi)(g(t)) \cdot g'(t)^2 \right].$$

\[10] In [34] the family of cylinder functions $\mathcal{X}C^\infty$ is introduced in [34, 7.24] and denoted by $\mathcal{Z}^\infty(\mathcal{P})$.

\[11] By a Cantor space we mean any non-empty totally disconnected perfect metrizable compact space.
The Malliavin–Shavgulidze image measure. Every $C^1$-function in $\mathcal{G}$ is a $C^1$-diffeomorphism of $S^1$, orientation-preserving since induced by a non-decreasing function, and every such diffeomorphism arises in this way. Furthermore, $\text{Diff}_+(S^1)$ embeds continuously into $\mathcal{G}$. It follows that $\mathcal{M}$ may be regarded as a (non-relabeled) measure on $\mathcal{G}$.

Example 4.18 (The Malliavin–Shavgulidze image measure). Consider the Borel probability measure $\mathcal{M}$ on $\mathcal{G}$. The measure $\mathcal{M}_1 := \text{pr}_{G^1}^\# \mathcal{M}$ is a well-defined Borel probability measure on $G^1$ by measurability (continuity) of $\text{pr}_{G^1}^\#$. The Malliavin–Shavgulidze image measure $\mathcal{S}$ is the Borel probability measure on $P$ defined by

$$\mathcal{S} := \chi^\#(\text{pr}_{G^1}^\# \mathcal{M}).$$

Assumptions $(\mathcal{P}_1)$ for $\mathcal{S}$ is readily verified from the properties of the Malliavian–Shavgulidze measure $\mathcal{M}$. In fact, $\mathcal{S}$ is concentrated on the set

$$\left( \text{Diff}_+(S^1)/\text{Isom_+}(S^1) \right) \cap \mathcal{M} \subset \mathcal{P}^m(S^1).$$

Assumption $(\mathcal{P}_5)$ is verified in Lemma 5.8 below, which suffices to establish assumption $(\mathcal{P})$ by Proposition 4.6.

Remark 4.19. Examples 4.11, 4.15 and 4.18 clarify that assumption $(\mathcal{P})$ poses no restriction to the subset of $\mathcal{P}$ where $\mathcal{P}$ is concentrated. Indeed, as argued above

- $N\mathcal{F}_{\lambda,\sigma}$-a.e. $\mu \in \mathcal{P}_2(S^1)$ is purely atomic;
- $\mathcal{P}_m$-a.e. $\mu \in \mathcal{P}_2(S^1)$ is singular continuous (w.r.t. the volume measure of $S^1$);
- $\mathcal{S}$-a.e. $\mu \in \mathcal{P}_2(S^1)$ is absolutely continuous (w.r.t. the volume measure of $S^1$).

Furthermore, it is readily seen that, if $\mathcal{P}$ and $\mathcal{P}'$ both satisfy assumption $(\mathcal{P})$, then so does any convex combination thereof. Thus, it is possible to construct a measure $\mathcal{P}$ on $\mathcal{P}_2(S^1)$ such that $\mathcal{P}$-a.e. $\mu$ has Lebesgue decomposition consisting of both a singular, a singular continuous and an absolutely continuous part.

5. Appendix.

5.1. On the notion of tangent bundle to $\mathcal{P}_2$. The concept of ‘tangent space’ to $\mathcal{P}_2$ at a point $\mu$ or ‘space of directions’ through $\mu$ has been widely investigated. (See [3, 20, 21, 22] and, especially, the bibliographical notes [2, §6.4].) At least the following three different notions are available

- the tangent space $T^\mu \mathcal{P}_2 := \text{cl}_{\mathcal{X}} \mathcal{X}^\mu$;
- the geometric tangent space, denoted here by $T_{\mu} \mathcal{P}_2$, defined in [21, 5.4];
- the pseudo-tangent space, denoted here by $T_{\mu}^{\text{Der}} \mathcal{P}_2 := \mathcal{X}_\mu$, considered as auxiliary space in [8, 20].

It was proven in [21, 6.1, 6.3] (cf. [2, 6.1]) that $T^\mu \mathcal{P}_2 \cong T_{\mu} \mathcal{P}_2$ if and only if $\mu \in \mathcal{P}_{\text{reg}}$; if otherwise, then $T^\mu \mathcal{P}_2$ embeds canonically non-surjectively in $T_{\mu} \mathcal{P}_2$ and the latter is not a Hilbert space. The relation between $T^\mu \mathcal{P}_2$ and $T_{\mu}^{\text{Der}} \mathcal{P}_2$ is made explicit in the following.
Preliminaries. In this section, we endow $\mathcal{C}^{\infty}(X)$ with its usual Fréchet\textsuperscript{12} topology $\tau_{\mathcal{C}^{\infty}(X)}$ and denote by $\mathcal{C}^{\infty}(X)^{*}$ the topological dual $(\mathcal{C}^{\infty}(X), \tau_{\mathcal{C}^{\infty}(X)})^{*}$ endowed with the weak* topology (see e.g. [38, §1.9]). Analogously, we endow $\mathcal{F}C^{\infty}$ with the locally convex metrizable linear topology $\tau_{\mathcal{F}C^{\infty}}$ induced by the countable family of semi-norms

$$|u|_{k} := \sup_{w_{1}, \ldots, w_{k} \in \mathcal{X}^{\infty}} \|\nabla_{w_{1}} \cdots \nabla_{w_{k}} u\|_{\mathcal{C}(\mathcal{P}_{2})}, \quad k \in \mathbb{N}_{0},$$

where it is understood that $|u|_{0}$ is but the uniform norm on $\mathcal{C}(\mathcal{P}_{2})$. We denote by $\mathcal{F}C^{\infty*}$ the topological dual of $(\mathcal{F}C^{\infty}, \tau_{\mathcal{F}C^{\infty}})$, endowed with the weak* topology.

Divergence operator (cf. [20, §2.3]). The divergence operator $\text{div}_{\mu} : \mathcal{X}^{\infty} \to \mathcal{C}^{\infty}(X)^{*}$ mapping

$$w \mapsto \langle \text{div}_{\mu} w \mid \cdot \rangle : f \mapsto - \int_{X} d\mu(x) (df(w))(x)$$

satisfies

$$\langle \text{div}_{\mu} w \mid f \rangle \leq \|\nabla f\|_{\mathcal{X}_{\mu}} \|w\|_{\mathcal{X}_{\mu}},$$

hence it extends by continuity to a (non-relabeled) operator $\text{div}_{\mu} : T^{\text{Der}}_{\mu} \mathcal{P}_{2} \to \mathcal{C}^{\infty}(X)^{*}$ and one has (see [20, 2.6])

$$T^{\text{Der}}_{\mu} \mathcal{P}_{2} = T^{\nabla}_{\mu} \mathcal{P}_{2} \oplus \ker \text{div}_{\mu},$$

where the symbol $\oplus$ denotes the orthogonal direct sum of Hilbert spaces.

On the one hand, it is clear that, if $\mu \in \mathcal{P}^{\infty, \infty}$, then $\ker \text{div}_{\mu}$ is non-trivial as soon as $\mathcal{X}^{\infty} \neq \mathcal{X}^{\infty}_{\infty}$. This holds in particular if $(X, g)$ has non-trivial de Rham cohomology group $H^{1}_{dR}(X; \mathbb{R})$. On the other hand (cf. [20, 2.8]), if $\eta \in \mathcal{P}$ has finite support, then

$$T^{\nabla}_{\eta} \mathcal{P}_{2} = T^{\text{Der}}_{\eta} \mathcal{P}_{2} = \bigoplus_{x \in \text{ptws } \eta} \left(T_{x}X, \eta_{x} \cdot g_{x}, \mathcal{X}_{\mu}\right).$$

Local derivations. Motivated by the definition, for finite-dimensional differential manifolds, of space of derivatives at a point (or pointwise derivations) (see e.g. [9, 2.22]), we define for fixed $\mu \in \mathcal{P}$ the linear functional $\partial_{w}^{\mu} : \mathcal{F}C^{\infty} \to \mathbb{R}$ by

$$\partial_{w}^{\mu} : u \mapsto \langle \nabla u(\mu) \mid w \rangle_{\mathcal{X}_{\mu}}.$$ 

Letting $\text{ev}_{\mu} : \mathcal{F}C^{\infty} \to \mathbb{R}$ be defined by $\text{ev}_{\mu}(u) := u(\mu)$, it is readily verified (cf. Lem 5.2 below) that $\partial_{w}^{\mu}$ satisfies Leibniz rule in the form

$$\partial_{w}^{\mu}(uv) = \text{ev}_{\mu}(v) \partial_{w}^{\mu} u + \text{ev}_{\mu}(u) \partial_{w}^{\mu} v.$$ 

We denote by $\text{Der}(\mathcal{F}C^{\infty})_{\mu} \subset \mathcal{F}C^{\infty*}$ the space of continuous linear functionals on $\mathcal{F}C^{\infty}$ satisfying (5.4), endowed with the trace topology. Since $\mathcal{F}C^{\infty*}$ is Hausdorff and complete, the (uniformly) continuous linear operator $\partial_{w}^{\mu} : \mathcal{X}^{\infty} \to \mathcal{F}C^{\infty*}$ extends to a uniquely defined non-relabeled operator $\partial_{w}^{\mu} : \mathcal{X}_{\mu} \to \mathcal{F}C^{\infty*}$ by [39, §1.5, Thm. 5.1]. Moreover,

$$\partial_{w}^{\mu}(u) \leq \|\nabla u(\mu)\|_{\mathcal{X}_{\mu}} \|w\|_{\mathcal{X}_{\mu}},$$

hence one has in fact $\partial_{w}^{\mu} \in \text{Der}(\mathcal{F}C^{\infty})_{\mu}$ for every $w \in \mathcal{X}_{\mu}$.

\textsuperscript{12}By a Fréchet space we mean a locally convex completely metrizable topological vector space.
PROPOSITION 5.1. Denote by \( j : f \mapsto f^{**} \) the canonical injection \( \mathcal{C}(X) \to \mathcal{C}(X)^{**} \).
Then, \( \text{div}_\mu(\cdot) = -\partial^\mu \circ j : \mathcal{X}_\mu \to \mathcal{C}^\infty(X)^* \) and \( \ker \text{div}_\mu = \ker \partial^\mu \subset \mathcal{X}_\mu \) as Hilbert spaces.

PROOF. For any \( f \in \mathcal{C}^\infty(X) \) and \( w \in \mathcal{X}_\infty \) it holds that

\[
(\delta^\mu_w \circ j)(f) = \delta^\mu_w(f^{**}) = \int_X d\mu(x) \langle \nabla f_x | w_x \rangle g = \mu(df(w)) = -\langle \text{div}_\mu w | f \rangle,
\]

that is \( \text{div}_\mu(\cdot)(-\nabla \cdot) = -\partial^\mu(j(-)) \) on \( \mathcal{X}_\infty \otimes \mathcal{C}^\infty(X) \). By (5.5) applied to \( u = f^{**} \in \mathcal{F}C^\infty \), the operator \( \partial^\mu \circ j : \mathcal{X}_\mu \to \mathcal{C}^\infty(X)^* \) may be extended to a uniquely defined non-relabeled operator \( \partial^\mu \circ j : \mathcal{X}_\mu \to \mathcal{C}^\infty(X)^* \) and the notation is consistent in the sense that this operator coincides with the previously defined extension of \( \partial^\mu \) applied to \( j \). Since both \( \partial^\mu \circ j \) and \( -\text{div}_\mu(\cdot) \) are linear and \( \| \cdot \|_{X_\mu} \)-continuous and coincide on the dense set \( \mathcal{X}_\infty \subset \mathcal{X}_\mu \), they coincide on the whole space \( \mathcal{X}_\mu \). It remains to show that \( \ker \partial^\mu = \ker \partial^\mu \circ j \), which follows immediately by noticing that for any \( u = F \circ f^{**} \in \mathcal{F}C^\infty \) and \( w \in \mathcal{X}_\mu \)

\[
\partial^\mu_w(u) = \sum_{i=0}^k (\partial_i F)(f^{**}) \int_X d\mu(x) \langle \nabla_x f_i | w_x \rangle g = -\sum_{i=0}^k (\partial_i F)(f^{**}) \langle \text{div}_\mu w | f_i \rangle
\]

This concludes the proof.

Tangent bundles. Let us denote by \( T^X \mathcal{P}_2 \) the tangent bundle to \( \mathcal{P}_2 \), set-wise defined as the disjoint union of \( T^X \mathcal{P}_2 \) varying \( \mu \in \mathcal{P}_2 \). The pseudo-tangent bundle \( T^\text{Der} \mathcal{P}_2 \) is analogously defined. Whereas this terminology is well-established, it is clear that \( T^X \mathcal{P}_2 \) is not a vector bundle in the standard sense — nor in any reasonable sense —, since it admits no local trivialization by reasons of the dimension of \( T^X \mathcal{P}_2 \). Indeed, for any \( x_0 \in X \) and every \( \varepsilon > 0 \) one can find a smooth function \( \rho_\varepsilon \in \mathcal{C}^\infty(X) \) such that \( \mu_\varepsilon := \rho_\varepsilon \mathbb{m} \in \mathcal{P}_m^\infty \) and \( W_2(\delta_{x_0}, \mu_\varepsilon) < \varepsilon \) yet \( T^{X}_{x_0} \mathcal{P}_2 \cong T_{x_0}X \) while \( T^X \mathcal{P}_2 \) is infinite-dimensional. The same is true for \( T^\text{Der} \mathcal{P}_2 \).

Despite this fact, the gradient \( \nabla u \) of a cylinder function \( u \in \mathcal{F}C^\infty \) may well be regarded as a ‘smooth section’ of \( T^X \mathcal{P}_2 \) since \( \nabla u(\mu) \in T^X \mathcal{P}_2 \) by (1.3). Again by (1.3) the space of all such gradients is a subspace of the space \( \mathcal{F}C^\infty \otimes_\mathbb{R} \mathcal{X}_\infty^\infty \) of \( \mathcal{F}C^\infty \)-linear combinations of gradient-type vector fields. This motivates the Definition 1.3 of cylinder vector fields \( \mathcal{X}C^\infty := \mathcal{F}C^\infty \otimes_\mathbb{R} \mathcal{X}_\infty^\infty \), henceforth regarded — in analogy to the case of finite-dimensional manifolds — as (a subspace of) the space of ‘smooth sections’ of the tangent bundle \( T^\text{Der} \mathcal{P}_2 \). In spite of Proposition 5.1, the fiber-bundle \( T^\text{Der} \mathcal{P}_2 \) does in fact convey more information than the fiber-bundle \( T^X \mathcal{P}_2 \).

Global derivations. Consider the space \( \text{Der}(\mathcal{F}C^\infty) \) of abstract \( \mathbb{R} \)-derivations of \( \mathcal{F}C^\infty \).

LEMMA 5.2. Let \( w \in \mathcal{X}_\infty \). Then, the map

\[
\delta^\mu_w : u \mapsto d\mu|_{t=0}(u \circ \Psi^{w,t}) = \langle \nabla u | w \rangle_X
\]

is an element of \( \text{Der}(\mathcal{F}C^\infty) \).

PROOF. One has

\[
d\mu|_{t=0}(\Psi^{w,t}) = \sum_{i=0}^k (\partial_i F)(f^{**}) \Psi^{w,0}_i \mu \times d\mu|_{t=0}(f^{**}) \Psi^{w,t}_i \mu = \sum_{i=0}^k (\partial_i F)(f^{**}) \times d\mu|_{t=0}(f_i \circ \Psi^{w,t})
\]
\[
\begin{align*}
&= \sum_{i}^{k} (\partial_i F)(f^{**}\mu) \times \mu \left( \left. d_t \right|_{t=0}(f_i \circ \psi^{w,t}) \right) = \sum_{i}^{k} (\partial_i F)(f^{**}\mu) \times \mu \langle \nabla f_i \mid w \rangle_{g} \\
&\quad \text{(5.7)} = \sum_{i}^{k} (\partial_i F)(f^{**}\mu) \times \langle \nabla f_i \mid w \rangle_{g}^{**} \mu \\\n&\quad \text{(5.8)} = \langle \nabla u(\mu) \mid w \rangle_{X_{\mu}}.
\end{align*}
\]

Since \( \langle \nabla f_i \mid w \rangle_{g} \in C^{\infty}(X) \) by the choice of \( f_i \) and \( w \), and since \( FC^{\infty} \) is an algebra, (5.7) shows that \( \partial_{w_\cdot} : FC^{\infty} \rightarrow FC^{\infty} \). The Leibniz rule is straightforward from the same property of \( d_t \), while \( FC^{\infty} \)-linearity is a consequence of the representation in (5.8).

\[\Box\]

**Proposition 5.3.** Let \( W \in \mathcal{X}C^{\infty} \) be as in (1.7). Then, the map

\[ \partial : W \rightarrow \partial W := \sum_{j}^{n} v_j \partial w_j \]

is a linear injection \( \partial : \mathcal{X}C^{\infty} \rightarrow \text{Der}(FC^{\infty}) \).

**Proof.** The fact that \( \partial W \in \text{Der}(FC^{\infty}) \) is a consequence of Lemma 5.2 and of the choice of the \( v_j \)'s. The \( FC^{\infty} \)-linearity is immediate, while the \( \mathcal{X}^{\infty} \)-linearity follows from (5.8).

Let now \( W \neq 0_{\mathcal{X}C^{\infty}} \), that is, there exists \( \mu_0 \in \mathcal{P} \) and \( x_0 \in X \) such that \( W(\mu_0)(x_0) \neq 0_{T_{x_0}X} \). Since \( W(\cdot)(x_0) \) is continuous and the set of purely atomic finitely supported probability measures is dense in \( \mathcal{P}_2 \) (see e.g. the proof of [42, 6.18]), we can find a purely atomic finitely supported \( \eta \in \mathcal{P} \) such that \( W(\eta)(x_0) \neq 0_{T_{x_0}X} \). Without loss of generality, up to choosing \( \eta' := \eta + \varepsilon \delta_{x_0} \) for some small \( \varepsilon > 0 \), we can assume \( \eta_{x_0} > 0 \) (for the notation see (2.1)). By standard arguments, there exists \( f \in C^{\infty}(X) \) such that \( \nabla f_{x_0} = W(\eta)(x_0) \). Moreover, since ptws \( \eta \) is discrete (finite), we can find \( g \in C^{\infty}(X) \) such that \( g \equiv 1 \) on an open neighborhood of \( x_0 \) and \( g \equiv 0 \) on an open neighborhood of every point in ptws \( \eta \) other than \( x_0 \). Set \( h = fg \) and notice that \( \nabla h_{x_0} = W(\eta)(x_0) \) while \( \nabla h = 0 \) for every point in ptws \( \eta \) other than \( x_0 \). Now,

\[ \partial W(h^{**})(\eta) = (\nabla h^{**} \mid W(\eta))_{X_{\eta}} = \int_{X} \text{d}\eta(x) \langle \nabla h_{x} \mid W(\eta)(x) \rangle_{g} = \eta_{x_0} |W(\eta)(x_0)|^{2}_{g} > 0. \]

Since \( \partial \) is linear, this shows that it is also injective, which concludes the proof. \( \Box \)

**Remark 5.4.** Proposition 5.3 is motivated by the analogy (see e.g. [9, 3.5.3]) with finite-dimensional compact differentiable manifolds, where the map

\[ \partial : \mathcal{X}^{\infty} \ni w \mapsto (\partial_{w_\cdot} : f \mapsto df(w)) \in \text{Der}(C^{\infty}(X)) \]

is straightforwardly injective, and surjective because of the classical Hadamard Lemma. In the case of \( \mathcal{P}_2 \), I do not know whether \( \partial \) is surjective, however, it should be noted that, in the case of infinite-dimensional smooth manifolds, this is not necessarily the case, again already at the pointwise scale (cf. e.g. [25, 28.7]).

Throughout all computations in Section 3, vector fields \( w \in \mathcal{X}^{\infty} \) ought to be interpreted as ‘smooth directions’ at every point \( \mu \in \mathcal{P} \). This is the right notion to be compared with the definition of directional derivative given in (1.4) in light of Proposition 5.3.
5.2. Auxiliary results on normalized mixed Poisson measures.

**Lemma 5.5.** The measure $\mathbb{P}$ defined in Example 4.11 satisfies assumption $(\text{P}_4)$.

**Proof.** Retain the notation in §4. By (4.1) and combining (4.2) with (4.3),
\[
d(\Phi_2\mathbb{P})(\mu) = d \left( \int_{\mathbb{R}^+} d\lambda(s) N_{\sigma} \Phi_2 \mathcal{P}_{s,\sigma} \right)(\mu) \\
= d \left( \int_{\mathbb{R}^+} d\lambda(s) \exp \left( s \cdot \sigma(1 - p_{\Phi}^{s,\sigma}) \right) N_1(R^{s,\sigma}_{\Phi} \cdot \mathcal{P}_{s,\sigma}) \right)(\mu).
\]

Noticing further that $N$ is injective on $\tilde{\Gamma}$ and denoting by $N^{-1}$ its right-inverse, the function $R_{\Phi}^{s,\sigma} \circ N^{-1}$ is well-defined on $\tilde{\Gamma}$, hence $\mathcal{P}_{s,\sigma}$-a.e. on $\Gamma$. It follows that
\[
d(\Phi_2\mathbb{P})(\mu) = \int_{\mathbb{R}^+} d\lambda(s) \exp \left( s \cdot \sigma(1 - p_{\Phi}^{s,\sigma}) \right) \cdot (R_{\Phi}^{s,\sigma} \circ N^{-1})(\mu) \cdot d(N_{\sigma} \mathcal{P}_{s,\sigma})(\mu).
\]
Moreover, $p_{\Phi}^{s,\sigma}$ for every $s > 0$ by definition (cf. (4.2)), thus $R_{\Phi}^{s,\sigma} = R_{\Phi}^{s,\sigma}$ and
\[
d(\Phi_2\mathbb{P})(\mu) = (R_{\Phi}^{s,\sigma} \circ N^{-1})(\mu) \int_{\mathbb{R}^+} d\lambda(s) \exp \left( s \cdot \sigma(1 - p_{\Phi}^{s,\sigma}) \right) \cdot d(N_{\sigma} \mathcal{P}_{s,\sigma})(\mu)
\]
(5.9)
where it is possible to pull $N_1$ outside the integral sign since the integrand does not depend on $\mu$.

Finally, for every measurable $A \subset \mathcal{P}$,
\[
e^{-c_{\lambda,\sigma,\Phi}}(R_{\Phi}^{s,\sigma} \circ N^{-1} \cdot \mathbb{P})A \leq (\Phi_2\mathbb{P})A \leq e^{c_{\lambda,\sigma,\Phi}}(R_{\Phi}^{s,\sigma} \circ N^{-1} \cdot \mathbb{P})A,
\]
where $c_{\lambda,\sigma,\Phi} := (\text{sup supp}\lambda) \sigma(1 - p_{\Phi}^{s,\sigma})$. Since $R_{\Phi}^{s,\sigma}(\gamma) > 0$ for $\mathcal{R}_{\lambda,\sigma}$-a.e. $\gamma \in \Gamma$, it follows from (5.10) that $\mathbb{P}$ and $\Phi_2\mathbb{P}$ are mutually absolutely continuous, hence the quasi-invariance assertion in $(\text{P}_4)$ holds. Letting $w \in \mathcal{X}^\infty$, equation (1.6) is similarly verified since $\#\text{supp}\mu < \infty$ for $\mathbb{P}$-a.e. $\mu$, hence, for all $t \in \mathbb{R}$,
\[
(R_{\Phi,w}^{s,\sigma} \circ N^{-1})(\mu) \geq \prod_{x \in \mu} \frac{(\psi_{w,t})^s \rho(x)}{\rho(x)} J_{\Phi,w,t}^m(x) \geq \left( \min_X \frac{(\psi_{w,t})^s \rho}{\rho} J_{\Phi,w,t}^m \right)^{\#\text{supp}\mu} > 0.
\]
This concludes the proof. \(\Box\)

**Lemma 5.6.** The measure $\mathbb{P}$ defined in Example 4.11 satisfies assumption $(\text{P}_3)$.

**Proof.** We show the assertion when $\lambda = \delta_1$, i.e. when $\mathbb{P} = N_2 \mathcal{P}_{\sigma}$, similarly to [1, Thm. 3.1]. The general case is readily proved by integration w.r.t. $\lambda$ in light of the mutual absolute continuity of $\mathcal{P}_{s,\sigma}$ w.r.t. $\mathcal{P}_\sigma$ (hence of their normalizations) for every choice of $s > 0$ (hence $\lambda$-a.e., cf. (4.4)).

**Preliminaries.** Retain the notation established in §4 and denote by $\beta^{\sigma} := \nabla \rho / \rho$ the logarithmic derivative of $\sigma$, which is well-defined on $X$ since $\rho \in C^1(X; \mathbb{R}^+)$. Let further $w \in \mathcal{X}^\infty$ and set
\[
\beta_{w}^{\sigma}(x) := \langle \beta_{x}^{\sigma} \mid w \rangle_{x} + \text{div}_m w_x,
\]
where $\text{div}_m$ denotes the divergence on $X$ with respect to the volume measure $m$. By integration by parts (cf. e.g. [1, (3.11)]) one can readily show that
\[
\nabla_{w}^{*} = -\nabla_{w} - \beta_{w}^{\sigma},
\]
where $\nabla_{w}^{*}$ denotes the adjoint of $\nabla_{w}$ in $L^2_0(X)$ and we denote the closure of $\nabla_{w}$ again by the same symbol.
Claim. Letting $B^\sigma_w := (\beta^\sigma_w)^*$, we claim that $\nabla_w := -\nabla_w - B^\sigma_w$ satisfies (1.5) for our choice of $\mathbb{P}$.

Some differentiation. For all $u \in \mathcal{FC}^\infty$ denote by the same symbol the natural extension of $u$ to $\mathcal{M}^\sigma_w$. By [1, Prop. 2.1] and several applications of (4.1) we have

$$ \int_{\mathcal{P}} dN_\sigma P \circ \Psi_w \cdot v = \int_{\mathcal{P}} dN_\sigma P \circ \Psi_w \cdot v \circ \nabla_w - B^\sigma_w. $$

Differentiating the l.h.s. of (5.11) under the sign of integral with respect to $t$ yields the l.h.s. of (1.5) by (1.4). Moreover, letting $\lambda = \delta_1$ in (5.9) yields

$$ \int_{\mathcal{P}} dN_\sigma P \circ \Psi_w \cdot v \circ \nabla_w - B^\sigma_w, $$

where $R^\sigma_{\Psi_w,t} \circ N^{-1}$ is well-defined as in Lemma 5.5. Also, with obvious meaning of the notation $x \in \mu$ for $\mu \in N(\Gamma)$,

$$ d_\mu(t = 0) \left( \exp (\sigma(1 - p_{\Psi_w,t}) \circ R^\sigma_{\Psi_w,t} \circ N^{-1}) \right) = \int_\chi d\sigma(x) \chi - \frac{\Psi_w \cdot \rho(x)}{\rho(x)} J^m_{\Psi_w,t}(x). $$

Now the arguments in the proof of [1, Thm. 3.1] apply verbatim, yielding

$$ d_\mu(t = 0) \left( \exp (\sigma(1 - p_{\Psi_w,t}) \circ R^\sigma_{\Psi_w,t} \circ N^{-1}) \right) = -B^\sigma_w. $$

Proof of the claim. Finally, differentiating $v \circ \nabla_w$ with respect to $t$ yields $-\nabla_w v$, again by (1.4). Combing this fact with (5.13), the derivative under integral sign with respect to $t$ of the r.h.s. of (5.12) reads $\int d\mathbb{P} u(-\nabla_w v - B^\sigma_w)$, which proves the claim. □

**Lemma 5.7.** The measure $\mathbb{P}$ defined in Example 4.11 satisfies assumption (P1).

**Proof.** By definition, $\mathbb{P}$ is concentrated on the set $N(\Gamma)$, which is dense in $\mathcal{P}$ (see e.g. [42, 6.18]). Let $U \neq \emptyset$ be open in $\mathcal{R}_2$. Then $U \cap N(\Gamma) \neq \emptyset$ by density of $N(\Gamma)$ in $\mathcal{R}_2$. By continuity of $N$ the set $\tilde{U} := N^{-1}(U \cap \Gamma) = N^{-1}(U \cap N(\Gamma)) \neq \emptyset$ is open in $\Gamma$. Since $\mathcal{R}_{\sigma,\lambda}$ is fully supported on $\Gamma$ (cf. [35, 5.6]), then $\mathbb{P} U = \mathcal{R}_{\sigma,\lambda} U > 0$. □

5.3. Auxiliary results on the Malliavin-Shavgulidze image measure.

**Lemma 5.8.** The measure $\mathcal{S}$ defined in Example 4.18 satisfies assumption (P5).

**Proof.** Retain all notation from §4.6. It follows from (4.8) that $(L_{\gamma})_a \mathcal{M} = \mathcal{M}$ for every $a \in \mathbb{S}^1$, hence $\mathcal{M}_1$ is quasi-invariant with respect to the left action of $\operatorname{Diff}_+^\infty(\mathbb{S}^1)$ (in fact: of $\operatorname{Diff}_+^2(\mathbb{S}^1)$) on $\mathcal{F}_1$ given by post-composition, i.e. (4.8) holds true with $\mathcal{M}_1$ in place of $\mathcal{M}$ for every Borel $A \subset \mathcal{F}_1$ and every $\phi$ in $\operatorname{Diff}_+^\infty(\mathbb{S}^1)$.

By definition of $\chi$ (eq. (4.7)), for every $\phi$ and $\Phi$ as in the beginning of §4, it holds that

$$ \chi(L_{\phi^{-1}}(g)) = (\phi^{-1} \circ g)_* \mathcal{M} = \phi^{-1}_* (g_* \mathcal{M}) = \phi^{-1}_* \chi(g) = \Phi^{-1}(\chi(g)). $$
As a consequence, for $\mu = \chi(g)$,

$$d\Phi_2^* S(\mu) = dS(\Phi^{-1}(\chi(g))) = dS(\chi(L_{\Phi^{-1}}(g))) = dM(1)(L_{\Phi^{-1}}(g)) = R_{\Phi}(g) \cdot dM(1)(g)$$

$$= (R_{\Phi} \circ \chi^{-1})(\mu) \cdot dS(\mu)$$

where $R_{\Phi}(g)$ is the Radon–Nikodým derivative

$$R_{\Phi}(g) := \exp \left[ \int_{S^1} dm(r) S(\phi^{-1})(g(r)) \cdot g'(r)^2 \right].$$

The conclusion straightforwardly follows from the form of $R_{\Phi}$. □

References.

[1] Albeverio, S., Kondratiev, Yu. G., and Röckner, M. Analysis and geometry on Configuration Spaces. J. Funct. Anal., 154:444–500, 1998.

[2] Ambrosio, L. and Gigli, N. A User’s Guide to Optimal Transport. In Ambrosio, L., Bressan, A., Helbing, D., Klar, A., and Zanella, E., editors, Modelling and Optimisation of Flows on Networks – Cetraro, Italy 2009, Editors: Benedetto Piccoli, Michel Rascle, volume 2062 of Lecture Notes in Mathematics, pages 1–155. Springer, 2013. Throughout the present work, we refer to (the numbering of) results in the extended version, available at http://cvgmt.sns.it/media/doc/paper/195/.

[3] Ambrosio, L., Gigli, N., and Savaré, G. Gradient Flows in Metric Spaces and in the Space of Probability Measures. Lectures in Mathematics - ETH Zurich. Birkhäuser, 2nd edition, 2008.

[4] Banyaga, A. The Structure of Classical Diffeomorphism Groups, volume 400 of Mathematics and Its Applications. Springer, 1997.

[5] Bogachev, V. I. and Mayer-Wolf, E. Some Remarks on Rademacher’s Theorem in Infinite Dimensions. Potential Anal., 5:23–30, 1996.

[6] Bouleau, N. and Hirsch, F. Dirichlet forms and analysis on Wiener space. De Gruyter, 1991.

[7] Chodosh, O. A lack of Ricci bounds for the entropic measure on Wasserstein space over the interval. J. Funct. Anal., 262(10):4570–4581, 2012.

[8] Chow, Y. T. and Gangbo, W. A partial Laplacian as an infinitesimal generator on the Wasserstein space. arXiv:1710.10536, Oct. 2017.

[9] Conlon, L. Differentiable Manifolds. Birkhäuser Advanced Texts. Birkhäuser, 2nd edition, 2001.

[10] De Philippis, G. and Rindler, F. On the Structure of $\mathcal{A}$-Free Measures and Applications. Ann. Math., 184:1017–1039, 2016.

[11] Dello Schiavo, L. The Dirichlet–Ferguson Diffusion on the Space of Probability Measures over a Closed Riemannian Manifold. Work in preparation.

[12] Dello Schiavo, L. Characteristic Functionals of Dirichlet Measures. arXiv:1810.09790, 2018.

[13] Dello Schiavo, L. and Lytvyov, E. A Mecke-type Characterization of the Dirichlet–Ferguson Measure. arXiv 1706.07602, 2017.

[14] Eberle, A. Girsanov-type transformations of local Dirichlet forms: an analytic approach. Osaka J. Math., 33(2):497–531, Jun 1996.

[15] Enchev, O. and Stroock, D. W. Rademacher’s theorem for Wiener functionals. Ann. Probab., 21(1):25–33, 1993.

[16] Ferguson, T. S. A Bayesian analysis of some nonparametric problems. Ann. Statist., pages 209–230, 1973.

[17] Figalli, A. Regularity of optimal transport maps (After Ma–Trudinger–Wang and Loeper). Astérisque J., 2010. Séminaire BOURBAKI 61ème année, 2008-2009, no 1009, Juin 2009.

[18] Frank, R. L., Lenz, D., and Wingert, D. Intrinsic metrics for non-local symmetric Dirichlet forms and applications to spectral theory. J. Funct. Anal., 266(8):4765–4808, 2014.

[19] Fukushima, M., Oshima, Y., and Takeda, M. Dirichlet forms and symmetric Markov processes, volume 19 of De Gruyter Studies in Mathematics. de Gruyter, extended edition, 2011.

[20] Gangbo, W., Kim, H. K., and Pacini, T. Differential Forms on Wasserstein Space and Infinite-dimensional Hamiltonian Systems. Mem. Am. Math. Soc., 211(995), 2010.

[21] Gigli, N. On the inverse implication of Brenier–McCann theorems and the structure of $(\mathcal{P}_2(M), W_2)$. Methods and Applications of Analysis, 18(2):127–158, 2011.

[22] Gigli, N. Second Order Analysis on $(\mathcal{P}_2(M), W_2)$. Mem. Am. Math. Soc., 216(1018), 2012.
[23] Grabowski, J. Free subgroups of diffeomorphism groups. *Fundam. Math.*, 131(2):103–121, 1988.

[24] Koskela, P. and Zhou, Y. Geometry and analysis of Dirichlet forms. *Adv. Math.*, 231:2755–2801, 2012.

[25] Kriegl, A. and Michor, P. W. *The Convenient Setting of Global Analysis*, volume 53 of *Mathematical Surveys and Monographs*. American Mathematical Society, 1997.

[26] Kuzmin, P. A. On circle diffeomorphisms with discontinuous derivatives and quasi-invariance subgroups of Malliavin–Shavgulidze measures. *J. Math. Anal. appl.*, 330:744–750, 2007.

[27] Lott, J. Some Geometric Calculations on Wasserstein Space. *Commun. Math. Phys.*, 277(2):423–437, 2007.

[28] Ma, Z.-M. and Röckner, M. *Introduction to the Theory of (Non-Symmetric) Dirichlet Forms*. Graduate Studies in Mathematics. Springer, 1992.

[29] Malliavin, P. M. and Malliavin, P. An Infinitesimally Quasi Invariant Measure on the Group of Diffeomorphisms of the Circle. In Kashiwara, M. and Miwa, T., editors, *ICM-90 Satellite Conference Proceedings – Special Functions*, pages 234–244. Springer, 1990.

[30] McShane, E. J. Extension of Range of Functions. *Bull. Amer. Math. Soc.*, 40(12):837–842, 1934.

[31] Nekvinda, A. and Zajíček, L. A simple proof of the Rademacher theorem. *Časopis pro Pěstování Matematiky*, 113(4):337–341, 1988.

[32] Otto, F. The Geometry of Dissipative Evolution Equations: The Porous Medium Equation. *Comm. Part. Diff. Eq.*, 26(1-2):101–174, 2001.

[33] Overbeck, L., Röckner, M., and Schmuland, B. An analytic approach to Fleming–Viot processes with interactive selection. *Ann. Probab.*, 23(1):1–36, 1995.

[34] Renesse, M.-K. von and Sturm, K.-T. Entropic measure and Wasserstein diffusion. *Ann. Probab.*, 37(3):1114–1191, 2009.

[35] Röckner, M. and Schied, A. Rademacher’s Theorem on Configuration Spaces and Applications. *J. Funct. Anal.*, 169:325–356, 1999.

[36] Sethuraman, J. A constructive definition of Dirichlet priors. *Stat. Sinica*, 4(2):639–650, 1994.

[37] Sturm, K.-T. Entropic Measure on Multidimensional Spaces. In Dalang, R., Dozzi, M., and Russo, F., editors, *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 of *Progress in Probability*, pages 261–277. Springer, 2011.

[38] Trèves, F. *Linear Partial Differential Equations with Constant Coefficients – Existence, Approximation and Regularity of Solutions*, volume 6 of *Mathematics and its Applications*. Gordon and Breach – Science Publishers, London, 1966.

[39] Trèves, F. *Topological Vector Spaces, Distributions and Kernels*, volume 25 of *Pure and Applied Mathematics*. Academic Press, New York, 1967.

[40] Tsilevich, N., Vershik, A. M., and Yor, M. An Infinite-Dimensional Analogue of the Lebesgue Measure and Distinguished Properties of the Gamma Process. *J. Funct. Anal.*, 185:274–296, 2001.

[41] Villani, C. *Topics in Optimal Transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[42] Villani, C. *Optimal transport, old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, 2009.

Institut für Angewandte Mathematik
Rheinische Friedrich-Wilhelms-Universität Bonn
Endenicher Aller 60
DE 53115 Bonn
Germany
E-mail: delloschiavo@iam.uni-bonn.de