DIFFERENTIAL FORMS ON WASSERSTEIN SPACE AND INFINITE-DIMENSIONAL HAMILTONIAN SYSTEMS

WILFRID GANGBO, HWA KIL KIM, AND TOMMASO PACINI

ABSTRACT. Let $\mathcal{M}$ denote the space of probability measures on $\mathbb{R}^D$ endowed with the Wasserstein metric. A differential calculus for a certain class of absolutely continuous curves in $\mathcal{M}$ was introduced in [5]. In this paper we develop a calculus for the corresponding class of differential forms on $\mathcal{M}$. In particular we prove an analogue of Green's theorem for 1-forms and show that the corresponding first cohomology group, in the sense of de Rham, vanishes. For $D = 2d$ we then define a symplectic distribution on $\mathcal{M}$ in terms of this calculus, thus obtaining a rigorous framework for the notion of Hamiltonian systems as introduced in [3]. Throughout the paper we emphasize the geometric viewpoint and the role played by certain diffeomorphism groups of $\mathbb{R}^D$.

1. INTRODUCTION

Historically speaking, the main goal of Symplectic Geometry has been to provide the mathematical formalism and the tools to define and study the most fundamental class of equations within classical Mechanics, *Hamiltonian ODEs*. Lie groups and group actions provide a key ingredient, in particular to describe the symmetries of the equations and to find the corresponding preserved quantities.

As the range of physical examples of interest expanded to encompass continuous media, fields, *etc.*, there arose the question of reaching an analogous theory for PDEs. It has long been understood that many PDEs should admit a reformulation as *infinite-dimensional Hamiltonian systems*. A deep early example of this is the work of Born-Infeld [9], [10] and Pauli [41], who started from a Hamiltonian formulation of Maxwell's equations to develop a quantum field theory in which the commutator of operators is analogous to the Poisson brackets used in the classical theory. Further examples include the wave and Klein-Gordon equations (cf. *e.g.* [15], [32]), the relativistic and non-relativistic Maxwell-Vlasov equations [8], [14], [33], and the Euler equations for incompressible fluids [7].

In each case it is necessary to define an appropriate phase space, build a symplectic or Poisson structure on it, find an appropriate energy functional, then show that the PDE coincides with the corresponding Hamiltonian flow. For various reasons, however, the results are often more formal than rigorous. In particular, existence and uniqueness theorems for PDEs require a good notion of weak solutions which need to be incorporated into the configuration and phase spaces; the geometric structure of these spaces needs to be carefully worked out; the functionals need the appropriate degree of regularity, *etc.* The necessary techniques can become quite complicated and *ad hoc*.

The purpose of this paper is to provide the basis for a new framework for defining and studying Hamiltonian PDEs. The configuration space we rely on is the *Wasserstein space*

$\mathcal{M}$ of non-negative Borel measures on $\mathbb{R}^D$ with total mass 1 and finite second moment. Over the past decade it has become clear that $\mathcal{M}$ provides a very useful space of weak solutions for those PDEs in which total mass is preserved. One of its main virtues is that it provides a unified theory for studying these equations. In particular, the foundation of the theory of Wasserstein spaces comes from Optimal Transport and Calculus of Variations, and these provide a toolbox which can be expected to be uniformly useful throughout the theory. Working in $\mathcal{M}$ also allows for extremely singular initial data, providing a bridge between PDEs and ODEs when the initial data is a Dirac measure.

The main geometric structure on $\mathcal{M}$ is that of a metric space. The geometric and analytic features of this structure have been intensively studied, cf. *e.g.* [5], [12], [13], [34], [40]. In particular the work [5] has developed a theory of *gradient flows* on metric spaces. In this work the technical basis for the notion of weak solutions to a flow on $\mathcal{M}$ is provided by the theory of *2-absolutely continuous curves*. In particular, [5] develops a differential calculus for this class of curves including a notion of "tangent space" for each $\mu \in \mathcal{M}$. Applied to $\mathcal{M}$, this allows for a rigorous reformulation of many standard PDEs as gradient flows on $\mathcal{M}$. Overall, this viewpoint has led to important new insights and results, cf. *e.g.* [2], [5], [13], [23], [40]. Topics such as geodesics, curvature and connections on $\mathcal{M}$ have also received much attention, cf. [4], [28], [29], [43], [44].

In the case $D = 2d$, recent work [3] indicates that other classes of PDEs can be viewed as *Hamiltonian flows* on $\mathcal{M}$. Developing this idea requires however a rigorous symplectic formalism for $\mathcal{M}$, adapted to the viewpoint of [5]. Our paper achieves two main goals. The first is to develop a general theory of differential forms on $\mathcal{M}$. We present this in Sections 4 and 5. This calculus should be thought of as dual to the calculus of absolutely continuous curves. Our main result here, Theorem 5.32, is an analogue of Green's theorem for 1-forms and leads to a proof that, in a specific sense, every closed 1-form on $\mathcal{M}$ is exact. The second goal is to show that there exists a natural symplectic and Hamiltonian formalism for $\mathcal{M}$ which is compatible with this calculus of curves and forms. The appropriate notions are defined and studied in Sections 6 and 7.

Given any mathematical construction, it is a fair question if it can be considered "the most natural" of its kind. It is well known for example that cotangent bundles admit a "canonical" symplectic structure. It is an important fact, discussed in Section 7, that on a non-technical level our symplectic formalism turns out to be formally equivalent to the Poisson structure considered in [33], cf. also [24] and [28]. From the geometric point of view it is clear that the structure in [33] is indeed an extremely natural choice. The choice of $\mathcal{M}$ as a configuration space is also both natural and classical. The difference between our paper and the previous literature appears precisely on the technical level, starting with the choice of geometric structure on $\mathcal{M}$. Specifically, whereas previous work tends to rely on various adaptations of differential geometric techniques, we choose the methods of Optimal Transport. The technical effort involved is justified by the final result: while previous studies are generally forced to restrict to smooth measures and functionals, our methods allow us to present a uniform theory which includes all singular measures and assumes very little regularity on the functionals. Sections 5.2 through 5.4 are an example of the technicalities this entails. Section 5.1 provides instead an example of the simplifications which occur when one assumes a higher degree of regularity.

By analogy with the case of gradient flows we expect that our framework and results will provide new impulse and direction to the development of the theory of Hamiltonian PDEs.

In particular, previous work and other work in progress inspired by these results lead to existence results for singular initial data [3], existence results for Hamiltonians satisfying weak regularity conditions [25], and to the development of a *weak KAM* theory for the nonlinear Vlasov equation [20]. It is to be expected that in the process of these developments our regularity assumptions will be even further relaxed so as to broaden the range of applications. We likewise expect that the geometric ideas underlying Symplectic Geometry and Geometric Mechanics will continue to play an important role in the development of the Wasserstein theory of Hamiltonian systems on $\mathcal{M}$. For example, in a very rough sense the relationship between our methods and those implicit in [33] can be thought of as analogous to the relationship between [18] and [7]. A connection between the choice of using Lie groups (as in [18] and [7]) or the space of measures as configuration spaces is provided by the process of *symplectic reduction*, cf. [31], [32].

It is an interesting question to what extent our results can be generalized to spaces of probability measures on other manifolds M. Regarding this issue, the situation is as follows. Many of the analytic foundations of our paper are provided by the work [5], which is based on the choice $M := \mathbb{R}^D$. In theory many results of [5] should be extendible to general Riemannian manifolds, but at present such an extension does not exist. Assuming that this extension will be obtained, we have written our paper in such a way as to make it clear how one might then try to extend our own results. This partly explains our emphasis on the geometric ideas and intuition underpinning our analytic definitions and results: exactly the same ideas would continue to hold for general manifolds M. Section 5.6 discusses how our results on cohomology depend on the choice $M := \mathbb{R}^D$. The situation regarding the symplectic structure is similar: one should expect most results to continue to hold for general symplectic manifolds M.

The above considerations make it worthwhile to stress the geometric viewpoint throughout this paper, with particular attention to the role played by certain group actions. It is important to emphasize, however, that we never try to use any form of infinite-dimensional geometry to prove our results. The reason behind this is that the various existing rigorous formulations of infinite-dimensional manifolds and Lie groups do not seem to be easily adaptable to our needs, cf. Section 3.4 for details. The typical approach adopted throughout our paper is thus as follows: (i) use geometric intuition to guide us towards specific choices of rigorous definitions, within the framework of [5]; (ii) prove theorems using the methods of [5] and Monge-Kantorovich theory; (iii) provide informal discussions of the geometric consequences of our results.

In recent years Wasserstein spaces have also been very useful in the field of Geometric Inequalities, cf. *e.g.* [1], [16], [17], [30]. Most recently, the theory of Wasserstein spaces has started producing results in Metric and Riemannian Geometry, cf. *e.g.* [29], [35], [43], [44]. Thus there exist at least three distinct communities which may be interested in these spaces: people working in Analysis/PDEs/Calculus of Variations, people in Geometrical Mechanics, people in Geometry. Concerning the exposition of our results, we have tried to take this into account in various ways: (i) by incorporating into the presentation an abundance of background material; (ii) by emphasizing the general geometric setting behind many of our constructions; (iii) by sometimes avoiding maximum generality in the results themselves. As much as possible we have also tried to keep the background material and the purely formal arguments separate from the main body of the article via a careful subdivision into sections and an appendix.

We now briefly summarize the contents of each section. Section 2 contains a brief introduction to the topological and differentiable structure (in the weak sense of [5]) of $\mathcal{M}$. Likewise, Appendix A reviews various notions from Differential Geometry including Lie derivatives, differential forms, Lie groups and group actions. The material in both is completely standard, but may still be useful to some readers. Section 3 provides a bridge between these two parts by revisiting the differentiable structure of $\mathcal{M}$ in terms of group actions. Although this point of view is maybe implicit in [5], it seems worthwhile to emphasize it. On a purely formal level, it leads to the conclusion that $\mathcal{M}$ should roughly be thought of as a *stratified* rather than a smooth manifold, see Section 3.2. It also relates the sets $\mathbb{R}^D \subset \mathcal{M} \subset (C_c^\infty)^*$. The first inclusion, based on Dirac measures, shows that the theory on $\mathcal{M}$ specializes by restriction to the standard theory on $\mathbb{R}^D$: this should be thought of as a fundamental test in this field, to be satisfied by any new theory on $\mathcal{M}$. The second inclusion provides background for relating the constructions of Section 6.2 to the work [33]. Overall, Section 3 is perhaps more intuitive than rigorous; however it does seem to offer a useful point of view on $\mathcal{M}$, providing intuition for the developments in the following sections. Section 4 defines the basic objects of study for a calculus on $\mathcal{M}$, namely differential forms, push-forward operations and an exterior differential operator. It also introduces the more general notion of *pseudo forms*. Pseudo forms are closely related to the group action: this is discussed in Section 4.3. Pseudo forms reappear in Section 5 as the main object of study, mainly because it seems both more natural and easier to control their regularity. The main result of this section is an analogue of Green's theorem for certain *annuli* in $\mathcal{M}$, Theorem 5.32. Stating and proving this result requires a good understanding of the differentiability and integrability properties of pseudo 1-forms. We achieve this in Sections 5.2 and 5.3. Our main application of Theorem 5.32 is Corollary 5.34, which shows that the 1-form defined by any closed regular pseudo 1-form on $\mathcal{M}$ is exact. Section 5.6 discusses the cohomological consequences of this result. In Section 6 we move on towards Symplectic Geometry, specializing to the case $D = 2d$. The main material is in Section 6.2: for each $\mu \in \mathcal{M}$ we introduce a particular subspace of the tangent space $T_\mu \mathcal{M}$ and show that it carries a natural symplectic structure. We also study the geometric properties of this symplectic distribution and define the notion of Hamiltonian systems on $\mathcal{M}$, thus providing a firm basis to the notion already introduced in [3]. Formally speaking, this distribution of subspaces is integrable and the above defines a *Poisson structure* on $\mathcal{M}$. The existence of a Poisson structure on $(C_c^\infty)^*$ had already been noticed in [33]: their construction is a formal infinite-dimensional analogue of Lie's construction of a canonical Poisson structure on the dual of any finite-dimensional Lie algebra. We review this construction in Section 7 and show that the corresponding 2-form restricts to ours on $\mathcal{M}$. In this sense our construction is formally equivalent to the Kirillov-Kostant-Souriau construction of a symplectic structure on the coadjoint orbits of the dual Lie algebra.

2. The topology on $\mathcal{M}$ and a differential calculus of curves

Let $\mathcal{M}$ denote the space of *Borel probability measures on $\mathbb{R}^D$ with bounded second moment*, i.e.

$$\mathcal{M} := \{\text{Borel measures on } \mathbb{R}^D : \mu \geq 0, \int_{\mathbb{R}^D} d\mu = 1, \int_{\mathbb{R}^D} |x|^2 \, d\mu < \infty\}.$$

The goal of this section is to show that $\mathcal{M}$ has a natural metric structure and to introduce a differential calculus due to [5] for a certain class of curves in $\mathcal{M}$. We refer to [5] and [45] for further details.

2.1. **The space of distributions.** Let C_c^∞ denote the space of compactly-supported smooth functions on $\mathbb{R}^D$. Recall that it admits the structure of a complete locally convex Hausdorff topological vector space, cf. *e.g.* [42] Section 6.2. Let $(C_c^\infty)^*$ denote the topological dual of C_c^∞, *i.e.* the vector space of continuous linear maps $C_c^\infty \to \mathbb{R}$. We endow $(C_c^\infty)^*$ with the *weak-* topology*, defined as the coarsest topology such that, for each $f \in C_c^\infty$, the induced evaluation maps
$$(C_c^\infty)^* \to \mathbb{R}, \quad \phi \mapsto \langle \phi, f \rangle$$
are continuous. In terms of sequences this implies that, $\forall f \in C_c^\infty$,
$$\phi_n \to \phi \Leftrightarrow \langle \phi_n, f \rangle \to \langle \phi, f \rangle.$$
Then $(C_c^\infty)^*$ is a locally convex Hausdorff topological vector space, cf. [42] Section 6.16. As such it has a natural differentiable structure.

The following fact may provide a useful context for the material of Section 2.2. We denote by $\mathcal{P}$ the set of all Borel probability measures on $\mathbb{R}^D$. A function f on $\mathbb{R}^D$ is said to be of *p-growth* (for some $p > 0$) if there exist constants $A, B \geq 0$ such that $|f(x)| \leq A + B|x|^p$. Let $C_b(\mathbb{R}^D)$ denote the set of continuous functions with 0-growth, *i.e.* the space of bounded continuous functions. As above we endow $(C_b(\mathbb{R}^D))^*$ with its natural weak-* topology, defined using test functions in $C_b(\mathbb{R}^D)$: this is also known as the *narrow* topology. Since $\mathcal{P}$ is contained in both $(C_b(\mathbb{R}^D))^*$ and $(C_c^\infty)^*$, it inherits two natural topologies. It is well known, cf. [5] Remarks 5.1.1 and 5.1.6, that the corresponding two notions of convergence of sequences coincide, but that the stronger topology induced from $(C_b(\mathbb{R}^D))^*$ is more interesting in that it is metrizable.

2.2. **The topology on $\mathcal{M}$.** Let $C_2(\mathbb{R}^D)$ denote the set of continuous functions with 2-growth, as in Section 2.1. We endow $(C_2(\mathbb{R}^D))^*$ with its natural weak-* topology, defined using test functions in $C_2(\mathbb{R}^D)$. As in Section 2.1, $\mathcal{M}$ is contained in both $(C_2(\mathbb{R}^D))^*$ and $(C_c^\infty)^*$. We will endow $\mathcal{M}$ with the topology induced from $(C_2(\mathbb{R}^D))^*$. Notice that $\mathcal{M}$ is a convex affine subset of $(C_2(\mathbb{R}^D))^*$. In particular it is contractible, so for $k \geq 1$ all its homology groups H_k and cohomology groups H^k, defined topologically, vanish. As in Section 2.1, it turns out that this topology is metrizable. A compatible metric can be defined as follows.

Definition 2.1. Let $\mu, \nu \in \mathcal{M}$. Consider

(2.1) $$W_2(\mu, \nu) := \left(\inf_{\gamma \in \Gamma(\mu, \nu)} \int_{\mathbb{R}^D \times \mathbb{R}^D} |x - y|^2 d\gamma(x, y) \right)^{1/2}.$$

Here, $\Gamma(\mu, \nu)$ denotes the set of Borel measures γ on $\mathbb{R}^D \times \mathbb{R}^D$ which have μ and ν as *marginals*, *i.e.* satisfying $\pi^1_\#(\gamma) = \mu$ and $\pi^2_\#(\gamma) = \nu$ where π^1 and π^2 denote the standard projections $\mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}^D$.

Equation 2.1 defines a distance on $\mathcal{M}$. It is known that the infimum in the right hand side of Equation 2.1 is always achieved. We will denote by $\Gamma_o(\mu, \nu)$ the set of γ which minimize this expression.

It can be shown that $(\mathcal{M}, W_2)$ is a separable complete metric space, cf. *e.g.* [5] Proposition 7.1.5. It is an important result from Monge-Kantorovich theory that

(2.2) $$W_2^2(\mu, \nu) = \sup_{u,v \in C(\mathbb{R}^D)} \left\{ \int_{\mathbb{R}^D} u \, d\mu + \int_{\mathbb{R}^D} v \, d\nu : u(x) + v(y) \leq |x - y|^2 \ \forall x, y \in \mathbb{R}^D \right\}.$$

Recall that μ is *absolutely continuous* with respect to Lebesgue measure $\mathcal{L}^D$, written $\mu << \mathcal{L}^D$, if it is of the form $\mu = \rho \mathcal{L}^D$ for some function $\rho \in L^1(\mathbb{R}^D)$. In this case for any $\nu \in \mathcal{M}$ there exists a unique map $T : \mathbb{R}^D \to \mathbb{R}^D$ such that $T_{\#}\mu = \nu$ and

$$(2.3) \qquad W_2^2(\mu, \nu) = \int_{\mathbb{R}^D} |x - T(x)|^2 d\mu(x),$$

cf. *e.g.* [5] or [19]. One refers to T as the *optimal map* that pushes μ forward to ν.

Example 2.2. Given $x \in \mathbb{R}^D$, let δ_x denote the corresponding *Dirac measure* on $\mathbb{R}^D$. Consider the set of such measures: this is a closed subset of $\mathcal{M}$ isometric to $\mathbb{R}^D$. More generally, let a_i $(i = 1, \ldots, n)$ be a fixed collection of distinct positive numbers such that $\sum a_i = 1$. Then the set of measures of the form $\sum a_i \delta_{x_i}$ constitutes a closed subset of $\mathcal{M}$, homeomorphic to $\mathbb{R}^{nD}$.

If $a_i \equiv 1/n$ then the set of measures of the form $\mu = \sum (1/n) \, \delta_{x_i}$ can be identified with $\mathbb{R}^{nD}$ quotiented by the set of permutations of n letters. This space is not a manifold in the usual sense; in the simplest case $D = 1$ and $n = 2$, it is homeomorphic to a closed half plane, which is a manifold with boundary.

Example 2.3. The set of all absolutely continuous measures is dense in $\mathcal{M}$. The set of all discrete measures, as in Example 2.2, is also dense in $\mathcal{M}$. Since these two sets are disjoint, neither is open nor closed in $\mathcal{M}$.

2.3. **Tangent spaces and the divergence operator.** Let $\mathcal{X}_c$ denote the space of compactly-supported smooth vector fields on $\mathbb{R}^D$. Set $\nabla C_c^\infty := \{\nabla f : f \in C_c^\infty\} \subset \mathcal{X}_c$. For $\mu \in \mathcal{M}$ let $L^2(\mu)$ denote the set of Borel maps $X : \mathbb{R}^D \to \mathbb{R}^D$ such that $||X||_\mu^2 := \int_{\mathbb{R}^D} |X|^2 d\mu$ is finite. Recall that $L^2(\mu)$ is a Hilbert space with the inner product

$$(2.4) \qquad \hat{G}_\mu(X, Y) := \int_{\mathbb{R}^D} \langle X, Y \rangle \, d\mu.$$

Remark 2.4. If $\mu = \rho \mathcal{L}^D$ for some $\rho : \mathbb{R}^d \to (0, \infty)$ such that $\int \rho dx = 1$ then the natural map $\mathcal{X}_c \to L^2(\mu)$ is injective. But in general it is not: for example if μ is the Dirac mass at x then two vector fields X, Y will be identified as soon as $X(x) = Y(x)$. However, the image of this map is always dense in $L^2(\mu)$.

In [5] Section 8.4, a "tangent space" is defined for each $\mu \in \mathcal{M}$ as follows.

Definition 2.5. Given $\mu \in \mathcal{M}$, let $T_\mu \mathcal{M}$ denote the closure of ∇C_c^∞ in $L^2(\mu)$. We call it the *tangent space* of $\mathcal{M}$ at μ. The *tangent bundle* $T\mathcal{M}$ is defined as the disjoint union of all $T_\mu \mathcal{M}$.

Definition 2.6. Given $\mu \in \mathcal{M}$ we define the *divergence operator*

$$div_\mu : \mathcal{X}_c \to (C_c^\infty)^*, \quad \langle div_\mu(X), f \rangle := - \int_{\mathbb{R}^D} df(X) \, d\mu.$$

Notice that the divergence operator is linear and that $\langle div_\mu(X), f \rangle \leq ||\nabla f||_\mu ||X||_\mu$. This proves that the operator div_μ extends to $L^2(\mu)$ by continuity; we will continue to use the same notation for the extended operator, so that $\mathrm{Ker}(div_\mu)$ is now a closed subspace of $L^2(\mu)$.

It follows from [5] Lemma 8.4.2 that, given any $\mu \in \mathcal{M}$, there is an orthogonal decomposition

$$(2.5) \qquad L^2(\mu) = \overline{\nabla C_c^\infty}^\mu \oplus \mathrm{Ker}(div_\mu).$$

We will denote by $\pi_\mu : L^2(\mu) \to \overline{\nabla C_c^\infty}^\mu$ the corresponding projection. Notice that each tangent space has a natural Hilbert space structure G_μ, obtained by restriction of $\hat{G}_\mu$ to $\overline{\nabla C_c^\infty}^\mu$.

Remark 2.7. Decomposition 2.5 shows that $T_\mu \mathcal{M}$ can also be identified with the quotient space $L^2(\mu)/\mathrm{Ker}(div_\mu)$: the map π_μ provides a Hilbert space isomorphism between these two spaces.

Example 2.8. Suppose that $x_1, \cdots, x_n$ are points in $\mathbb{R}^D$ and $\mu = \sum_{i=1}^n (1/n) \delta_{x_i}$. Fix $\xi \in L^2(\mu)$. Set $4r := \min_{x_i \neq x_j} |x_i - x_j|$ and define

$$(2.6) \qquad \varphi(x) = \begin{cases} \langle x, \xi(x_i) \rangle & \text{if} \quad x \in B_{2r}(x_i) \quad i = 1, \cdots, n \\ 0 & \text{if} \quad x \notin \cup_{i=1}^n B_{2r}(x_i). \end{cases}$$

Let $\eta \in C_c^\infty$ be a symmetric function such that $\int_{\mathbb{R}^D} \eta dx = 1$, $\eta \geq 0$ and η is supported in the closure of $B_r(0)$. Then $\bar{\varphi} := \eta * \varphi \in C_c^\infty$ and $\nabla \bar{\varphi}$ coincides with ξ on $\cup_{i=1}^n B_r(x_i)$. Consequently, $L^2(\mu) = T_\mu \mathcal{M}$ and $\mathrm{Ker}(div_\mu) = \{0\}$. In particular if the points x_i are distinct then $L^2(\mu)$ can be identified with $\mathbb{R}^{nD}$. If on the other hand all the points coincide, *i.e.* $x_i \equiv x$, then $\mu = \delta_x$ and $L^2(\mu) \simeq \mathbb{R}^D$.

Consider for example the simplest case $D = 1$, $n = 2$. As seen in Example 2.2 the corresponding space of Dirac measures is homeomorphic to a closed half plane. We now see that at any interior point, corresponding to $x_1 \neq x_2$, the tangent space is $\mathbb{R}^2$. At any boundary point, corresponding to $x_1 = x_2$, the tangent space is $\mathbb{R}$. One should compare this with the usual differential-geometric definition of tangent planes on a manifold with boundary, cf. *e.g.* [21]: in that case the tangent plane at a boundary point would be $\mathbb{R}^2$. We will come back to this in Section 3.2.

Remark 2.9. Decomposition 2.5 extends the standard orthogonal Hodge decomposition of a smooth L^2 vector field X on $\mathbb{R}^D$:

$$X = \nabla u + X',$$

where u is defined as the unique smooth solution in $W^{1,2}$ of $\Delta u = div(X)$ and $X' := X - \nabla u$.

In particular, Decomposition 2.5 shows that $\overline{\nabla C_c^\infty}^\mu \cap \mathrm{Ker}(div_\mu) = \{0\}$. The analogous statement with respect to the measure $\mathcal{L}^D$ is that the only harmonic function on $\mathbb{R}^D$ in $W^{1,2}$ is the function $u \equiv 0$.

2.4. **Analytic justification for the tangent spaces.** Following [5] we now provide an analytic justification for the above definition of tangent spaces for $\mathcal{M}$. A more geometric justification, using group actions, will be given in Section 3.2.

Suppose we are given a curve $\sigma : (a, b) \to \mathcal{M}$ and a Borel vector field $X : (a, b) \times \mathbb{R}^D \to \mathbb{R}^D$ such that $X_t \in L^2(\sigma_t)$. Here, we have written σ_t in place of $\sigma(t)$ and X_t in place of $X(t)$. We will write

$$(2.7) \qquad \frac{\partial \sigma}{\partial t} + div_\sigma(X) = 0$$

if the following condition holds: for all $\phi \in C_c^\infty((a,b) \times \mathbb{R}^D)$,

$$(2.8) \qquad \int_a^b \int_{\mathbb{R}^D} \Big(\frac{\partial \phi}{\partial t} + d\phi(X_t) \Big) \, d\sigma_t \, dt = 0,$$

i.e. if Equation 2.7 holds in the sense of distributions. Given σ_t, notice that if Equation 2.7 holds for X then it holds for $X + W$, for any Borel map $W : (a,b) \times \mathbb{R}^D \to \mathbb{R}^D$ such that $W_t \in \mathrm{Ker}(div_{\sigma_t})$.

The following definition and remark can be found in [5] Chapter 1.

Definition 2.10. Let $(\mathbb{S}, \mathrm{dist})$ be a complete metric space. A curve $t \in (a,b) \mapsto \sigma_t \in \mathbb{S}$ is *2-absolutely continuous* if there exists $\beta \in L^2(a,b)$ such that $\mathrm{dist}(\sigma_t, \sigma_s) \leq \int_s^t \beta(\tau)d\tau$ for all $a < s < t < b$. We then write $\sigma \in AC_2(a,b;\mathbb{S})$. For such curves the limit $|\sigma'|(t) := \lim_{s \to t} \mathrm{dist}(\sigma_t, \sigma_s)/|t-s|$ exists for $\mathcal{L}^1$-almost every $t \in (a,b)$. We call this limit the *metric derivative* of σ at t. It satisfies $|\sigma'| \leq \beta$ $\mathcal{L}^1$-almost everywhere.

Remark 2.11. (i) If $\sigma \in AC_2(a,b;\mathbb{S})$ then $|\sigma'| \in L^2(a,b)$ and $\mathrm{dist}(\sigma_s, \sigma_t) \leq \int_s^t |\sigma'|(\tau)d\tau$ for $a < s < t < b$. We can apply Hölder's inequality to conclude that $\mathrm{dist}^2(\sigma_s, \sigma_t) \leq c|t-s|$ where $c = \int_a^b |\sigma'|^2(\tau)d\tau$.

(ii) It follows from (i) that $\{\sigma_t|\, t \in [a,b]\}$ is a compact set, so it is bounded. For instance, given $x \in \mathbb{S}$, the triangle inequality proves that $\mathrm{dist}(\sigma_s, x) \leq \sqrt{c|s-a|} + \mathrm{dist}(\sigma_a, x)$.

We now recall [5] Theorem 8.3.1. It shows that the definition of tangent space given above is flexible enough to include the velocities of any "good" curve in $\mathcal{M}$.

Proposition 2.12. *If $\sigma \in AC_2(a,b;\mathcal{M})$ then there exists a Borel map $v : (a,b) \times \mathbb{R}^D \to \mathbb{R}^D$ such that $\frac{\partial\sigma}{\partial t} + div_\sigma(v) = 0$ and $v_t \in L^2(\sigma_t)$ for $\mathcal{L}^1$-almost every $t \in (a,b)$. We call v a velocity for σ. If w is another velocity for σ then the projections $\pi_{\sigma_t}(v_t)$, $\pi_{\sigma_t}(w_t)$ coincide for $\mathcal{L}^1$-almost every $t \in (a,b)$. One can choose v such that $v_t \in \overline{\nabla C_c^\infty}^{\sigma_t}$ and $||v_t||_{\sigma_t} = |\sigma'|(t)$ for $\mathcal{L}^1$-almost every $t \in (a,b)$. In that case, for $\mathcal{L}^1$-almost every $t \in (a,b)$, v_t is uniquely determined. We denote this velocity $\dot\sigma$ and refer to it as the velocity of minimal norm, since if w_t is any other velocity associated to σ then $||\dot\sigma_t||_{\sigma_t} \leq ||w_t||_{\sigma_t}$ for $\mathcal{L}^1$-almost every $t \in (a,b)$.*

The following remark can be found in [5] Lemma 1.1.4 in a more general context.

Remark 2.13 (Lipschitz reparametrization). Let $\sigma \in AC_2(a,b;\mathcal{M})$ and v be a velocity associated to σ. Fix $\alpha > 0$ and define $S(t) = \int_a^t (\alpha + ||v_\tau||_{\sigma_\tau})d\tau$. Then $S : [a,b] \to [0,L]$ is absolutely continuous and increasing, with $L = S(b)$. The inverse of S is a function whose Lipschitz constant is less than or equal to $1/\alpha$. Define

$$\bar\sigma_s := \sigma_{S^{-1}(s)}, \qquad \bar v_s := \dot S^{-1}(s)v_{S^{-1}(s)}.$$

One can check that $\bar\sigma \in AC_2(0,L;\mathcal{M})$ and that $\bar v$ is a velocity associated to $\bar\sigma$. Fix $t \in (a,b)$ and set $s := S(t)$. Then $v_t = \dot S(t)\bar v_{S(t)}$ and $||\bar v_s||_{\bar\sigma_s} = \frac{||v_t||_{\sigma_t}}{\alpha + ||v_t||_{\sigma_t}} < 1$.

3. The calculus of curves, revisited

The goal of this section is to revisit the material of Section 2 from a more geometric viewpoint. Many of the results presented here are purely formal, but they may provide some insight into the structure of $\mathcal{M}$. They also provide useful intuition into the more rigorous results contained in the sections which follow. We refer to Appendix A for notation and terminology.

3.1. **Embedding the geometry of $\mathbb{R}^D$ into $\mathcal{M}$.** We have already seen in Example 2.2 that Dirac measures provide a continuous embedding of $\mathbb{R}^D$ into $\mathcal{M}$. Many aspects of the standard geometry of $\mathbb{R}^D$ can be recovered inside $\mathcal{M}$, and various techniques which we will be using for $\mathcal{M}$ can be seen as an extension of standard techniques used for $\mathbb{R}^D$.

One example of this is provided by Example 2.8, which shows that the standard notion of tangent space on $\mathbb{R}^D$ coincides with the notion of tangent spaces on $\mathcal{M}$ introduced by [5].

Another simple example concerns calculus on $\mathbb{R}^D$, as follows. Consider the space of *volume forms* on $\mathbb{R}^D$, *i.e.* the smooth never-vanishing D-forms. Under appropriate normalization and decay conditions, these define a subset of $\mathcal{M}$. Given a vector field $X \in \mathcal{X}_c$ and a volume form α, there is a standard geometric definition of $div_\alpha(X)$ in terms of Lie derivatives: namely, $\mathcal{L}_X \alpha$ is also a D-form so we can define $div_\alpha(X)$ to be the unique smooth function on $\mathbb{R}^D$ such that

$$(3.1) \qquad div_\alpha(X)\alpha = \mathcal{L}_X \alpha.$$

In particular, it is clear from this definition and Lemma A.3 that $X \in \text{Ker}(div_\alpha)$ iff the corresponding flow preserves the volume form.

Cartan's formula A.13 together with Green's theorem for $\mathbb{R}^D$ shows that div_α is the negative formal adjoint of d with respect to α, *i.e.*

$$\int_{\mathbb{R}^D} f \, div_\alpha(X)\alpha = -\int_{\mathbb{R}^D} df(X) \, \alpha, \quad \forall f \in C_c^\infty.$$

In particular, $div_\alpha(X)\alpha$ satisfies Equation 2.6. In this sense Definition 2.6 extends the standard geometric definition of divergence to the whole of $\mathcal{M}$.

3.2. **The intrinsic geometry of $\mathcal{M}$.** It is appealing to think that, in some weak sense, the results of Section 2.4 can be viewed as a way of using the Wasserstein distance to describe an "intrinsic" differentiable structure on $\mathcal{M}$. This structure can be alternatively viewed as follows.

Let $\phi : \mathbb{R}^D \to \mathbb{R}^D$ be a Borel map and $\mu \in \mathcal{M}$. Recall that the *push-forward* measure $\phi_{\#}\mu \in \mathcal{M}$ is defined by setting $\phi_{\#}\mu(A) := \mu(\phi^{-1}(A))$, for any open subset $A \subseteq \mathbb{R}^D$. Let $\text{Diff}_c(\mathbb{R}^D)$ denote the Id-component of the Lie group of diffeomorphisms of $\mathbb{R}^D$ with compact support, cf. Section A.4. Choose any $X \in \mathcal{X}_c$ and let ϕ_t denote the flow of X. Given any $\mu \in \mathcal{M}$, it is simple to verify that $\mu_t := \phi_{t\#}\mu$ is a path in $\mathcal{M}$ with velocity X in the sense of Proposition 2.12. Notice that in this case the velocity is defined for all t, rather than only for almost every t. In particular the minimal velocity of μ_t at $t = 0$ is $\pi_\mu(X) \in T_\mu \mathcal{M}$. From the point of view of Section A.2, this construction can be rephrased as follows. The map

$$(3.2) \qquad \text{Diff}_c(\mathbb{R}^D) \times \mathcal{M} \to \mathcal{M}, \quad (\phi, \mu) \mapsto \phi_{\#}\mu$$

is continuous and defines a *left action* of $\text{Diff}_c(\mathbb{R}^D)$ on $\mathcal{M}$. The map

$$\mathcal{M} \to T\mathcal{M}, \quad \mu \to \pi_\mu(X) \in T_\mu \mathcal{M}$$

then defines the *fundamental vector field* associated to X in the sense of Section A.2.

According to Section A.2, the orbit and stabilizer of any fixed $\mu \in \mathcal{M}$ are:

$$\mathcal{O}_\mu := \{\nu \in \mathcal{M} : \nu = \phi_{\#}\mu, \text{ for some } \phi \in \text{Diff}_c(\mathbb{R}^D)\},$$

$$\text{Diff}_{c,\mu}(\mathbb{R}^D) := \{\phi \in \text{Diff}_c(\mathbb{R}^D) : \phi_{\#}\mu = \mu\}.$$

Formally, $\mathrm{Diff}_{c,\mu}(\mathbb{R}^D)$ is a Lie subgroup of $\mathrm{Diff}_c(\mathbb{R}^D)$ and $\mathrm{Ker}(div_\mu)$ is its Lie algebra. The map

$$j : \mathrm{Diff}_c(\mathbb{R}^D)/\mathrm{Diff}_{c,\mu}(\mathbb{R}^D) \to \mathcal{O}_\mu, \quad [\phi] \mapsto \phi_\# \mu$$

defines a 1:1 relationship between the quotient space and the orbit of μ. Lemma A.15 suggests that $\mathcal{O}_\mu$ is a smooth manifold inside the topological space $\mathcal{M}$ and that the isomorphism $\nabla j : \mathcal{X}_c/\mathrm{Ker}(div_\mu) \to T_\mu \mathcal{O}_\mu$ coincides with the map determined by the construction of fundamental vector fields. Notice that, up to L^2_μ-closure, the space $\mathcal{X}_c/\mathrm{Ker}(div_\mu)$ is exactly the space introduced in Definition 2.5. This indicates that the tangent spaces of Section 2.3 should be thought of as "tangent" not to the whole of $\mathcal{M}$, but only to the leaves of the foliation induced by the action of $\mathrm{Diff}_c(\mathbb{R}^D)$. In other words $\mathcal{M}$ should be thought of as a *stratified manifold, i.e.* as a topological space with a foliation and a differentiable structure defined only on each leaf of the foliation. This point of view is purely formal but it corresponds exactly to the situation already described for Dirac measures, cf. Example 2.8.

Recall from Proposition 2.12 the relationship between the class of 2-absolutely continuous curves and these tangent spaces. This result can be viewed as the expression of a strong compatibility between two natural but *a priori* distinct structures on $\mathcal{M}$: the Wasserstein topology and the group action.

Remark 3.1. The claim that the Lie algebra of $\mathrm{Diff}_{c,\mu}(\mathbb{R}^D)$ is $\mathrm{Ker}(div_\mu)$ can be supported in various ways. For example, assume ϕ_t is a curve of diffeomorphisms in $\mathrm{Diff}_{c,\mu}(\mathbb{R}^D)$ and that X_t satisfies Equation A.8. The following calculation is the weak analogue of Lemma A.3. It shows that $X_t \in \mathrm{Ker}(div_\mu)$:

$$\int df(X_t)\, d\mu = \int df(X_t)\, d(\phi_{t\#}\mu) = \int df_{|\phi_t}(X_{t|\phi_t})\, d\mu = \int d/dt(f \circ \phi_t)\, d\mu$$

$$= d/dt \int f \circ \phi_t\, d\mu = d/dt \int f\, d(\phi_{t\#}\mu) = d/dt \int f\, d\mu$$

$$= 0.$$

It is also simple to check that $\mathrm{Ker}(div_\mu)$ is a Lie subalgebra of $\mathcal{X}_c$, *i.e.* if $X, Y \in \mathrm{Ker}(div_\mu)$ then $[X, Y] \in \mathrm{Ker}(div_\mu)$. To show this, let $f \in C^\infty_c$. Then:

$$\langle div_\mu([X, Y]), f \rangle = -\int_{\mathbb{R}^D} df([X, Y])\, d\mu = -\int_{\mathbb{R}^d} dg(X)\, d\mu + \int_{\mathbb{R}^d} dh(Y)\, d\mu$$

$$= \langle div_\mu(X), g \rangle - \langle div_\mu(Y), h \rangle$$

$$= 0,$$

where $g := df(Y)$ and $h := df(X)$.

Finally, assume μ is a smooth volume form on a compact manifold M. In this situation Hamilton [22] proved that $\mathrm{Diff}_\mu(M)$ is a Fréchet Lie subgroup of $\mathrm{Diff}(M)$ and that the Lie algebra of $\mathrm{Diff}_\mu(M)$ is the space of vector fields $X \in \mathcal{X}(M)$ satisfying the condition $\mathcal{L}_X \mu = 0$. As seen in Section 3.1 this space coincides with $\mathrm{Ker}(div_\mu)$.

3.3. **Embedding the geometry of $\mathcal{M}$ into $(C^\infty_c)^*$.** We can also view $\mathcal{M}$ as a subspace of $(C^\infty_c)^*$. It is then interesting to compare the corresponding geometries, as follows.

Consider the natural left action of $\mathrm{Diff}_c(\mathbb{R}^D)$ on $\mathbb{R}^D$ given by $\phi \cdot x := \phi(x)$. As in Section A.2, this induces a left action on the spaces of forms Λ^k and in particular on the space of

functions $C_c^\infty = \Lambda^0$, as follows:

$$\mathrm{Diff}_c(\mathbb{R}^D) \times C_c^\infty \to C_c^\infty, \quad \phi \cdot f := (\phi^{-1})^* f = f \circ \phi^{-1}.$$

By duality there is an induced left action on the space of distributions given by

$$\mathrm{Diff}_c(\mathbb{R}^D) \times (C_c^\infty)^* \to (C_c^\infty)^*, \quad \langle (\phi \cdot \mu), f \rangle := \langle \mu, (\phi^{-1} \cdot f) \rangle = \langle \mu, (f \circ \phi) \rangle.$$

Notice that we have introduced inverses to ensure that these are left actions, cf. Remark A.9. It is clear that this extends the action already defined in Section 3.2 on the subset $\mathcal{M} \subset (C_c^\infty)^*$. In other words, the natural immersion $i : \mathcal{M} \to (C_c^\infty)^*$ is *equivariant* with respect to the action of $\mathrm{Diff}_c(\mathbb{R}^D)$, *i.e.* $i(\phi_\# \mu) = \phi \cdot i(\mu)$.

As mentioned in Section 2.1, $(C_c^\infty)^*$ has a natural differentiable structure. In particular it has well-defined tangent spaces $T_\mu(C_c^\infty)^* = (C_c^\infty)^*$. For each $\mu \in \mathcal{M}$, using the notation of Section 3.2, composition gives an immersion

$$i \circ j : \mathrm{Diff}_c(\mathbb{R}^D)/\mathrm{Diff}_{c,\mu}(\mathbb{R}^D) \to \mathcal{O}_\mu \to (C_c^\infty)^*.$$

This induces an injection between the corresponding tangent spaces

$$\nabla(i \circ j) : \mathcal{X}_c/\mathrm{Ker}(div_\mu) \to T_\mu(C_c^\infty)^*.$$

Notice that, using the equivariance of i,

$$\langle \nabla(i \circ j)(X), f \rangle = \langle \nabla i(d/dt(\phi_{t\#}\mu)_{|t=0}), f \rangle = \langle d/dt(i(\phi_{t\#}\mu))_{|t=0}, f \rangle = \langle d/dt(\phi_t \cdot \mu)_{|t=0}, f \rangle$$
$$= d/dt \langle \mu, f \circ \phi_t \rangle_{|t=0} = \langle \mu, d/dt(f \circ \phi_t)_{|t=0} \rangle = \langle \mu, df(X) \rangle$$
$$= -\langle div_\mu(X), f \rangle.$$

In other words, the negative divergence operator can be interpreted as the natural identification between $T_\mu\mathcal{M}$ and the appropriate subspace of $(C_c^\infty)^*$.

More generally, we can compare the calculus of curves in $\mathcal{M}$ with the calculus of the corresponding curves in $(C_c^\infty)^*$. Given any sufficiently regular curve of distributions $t \to \mu_t \in (C_c^\infty)^*$, we can define tangent vectors $\tau_t := \lim_{h \to 0} \frac{\mu_{t+h} - \mu_t}{h} \in T_{\mu_t}(C_c^\infty)^*$. Assume that μ_t is strongly continuous, in the sense that the evaluation map

$$(a,b) \times C_c^\infty \to \mathbb{R}, \quad (t, f) \mapsto \langle \mu_t, f \rangle$$

is continuous. Notice that $\mu = \mu_t$ defines a distribution on the product space $(a,b) \times \mathbb{R}^D$: $\forall f = f_t(x) \in C_c^\infty((a,b) \times \mathbb{R}^D)$,

$$\langle \mu, f \rangle := \int_a^b \langle \mu_t, f_t \rangle \, dt.$$

One can check that $\frac{d}{dt}\langle \mu_t, f_t \rangle = \langle \tau_t, f_t \rangle + \langle \mu_t, \frac{\partial f_t}{\partial t} \rangle$, so

$$(3.3) \qquad \int_a^b \langle \mu_t, \frac{\partial f_t}{\partial t} \rangle + \langle \tau_t, f_t \rangle \, dt = 0.$$

Equation 3.3 shows that if $\mu_t \in \mathcal{M}$ and $\tau_t = -div_{\mu_t}(X_t)$ then μ_t satisfies Equation 2.8. In other words, the defining equation for the calculus on $\mathcal{M}$, Equation 2.7, is the natural weak analogue of the statement $\lim_{h \to 0} \frac{\mu_{t+h} - \mu_t}{h} = -div_{\mu_t}(X_t)$.

Roughly speaking, the content of Proposition 2.12 is that if $\mu_t \in \mathcal{M}$ is 2-absolutely continuous then, for almost every t, τ_t exists and can be written as $-div_{\mu_t}(X_t)$ for some t-dependent vector field X_t on $\mathbb{R}^D$.

Remark 3.2. One should think of Equation 2.7, *i.e.* $d/dt(\mu_t) = -div_{\mu_t}(X_t)$, as an ODE on the submanifold $\mathcal{M} \subset (C_c^\infty)^*$ rather than on the abstract manifold $\mathcal{M}$, in the sense that the right hand side is an element of $T_{\mu_t}(C_c^\infty)^*$ rather than an element of $T_{\mu_t}\mathcal{M}$. Using $\nabla(i \circ j)^{-1}$ we can rewrite this equation as an ODE on the abstract manifold $\mathcal{M}$, *i.e.* $d/dt(\mu_t) = \pi_{\mu_t}(X)$.

3.4. **Further comments.** There exists an extensive literature concerning how to make infinite-dimensional geometry rigorous. The first step is to provide rigorous definitions of infinite-dimensional manifolds and of infinite-dimensional Lie groups. The works [18], [27], [37] and [38] are examples of standard references in this field. In all these cases the starting point is a notion of manifold built by gluing together charts which are open subsets of locally convex vector spaces (plus some completeness condition). These references are often useful when one wants to make a "formal" study of PDE rigorous; in particular, when the space of solutions is, in some sense, an infinite-dimensional Lie group (as in [7]).

In this paper, however, we do not rely on the above frameworks. The main reason is, quite simply, the fact that the "differentiable structure" on $\mathcal{M}$ introduced by [5] is not based on the above notion of manifold: as discussed in Section 3.2, it uses a much weaker notion and none of the results presented in this paper require anything more than this. It should also be emphasized that the relationship, discussed in Chapters 2 and 3, between the Wasserstein metric on $\mathcal{M}$, group actions and the theory of [5] shows that the latter is extremely natural within this context.

Another reason is that we want to keep regularity assumptions to a minimum. In particular, we want to avoid making unnecessary restrictions on the smoothness of measures (required by [18]) and of velocity fields.

It may also be worthwhile to mention that it is not clear if the above references lead to a general theory of "infinite-dimensional homogeneous spaces", which is in some sense the geometry underlying this paper. Specifically, Section 3.2 introduces the idea that $\mathcal{M}$ is a "stratified manifold" and that each stratum is the orbit of a certain group action. These orbits are of the form G/H, where G is the infinite-dimensional Lie group of diffeomorphisms (as in the above references) and H is the subgroup of diffeomorphisms which preserve a given measure. Thus our space of solutions $\mathcal{M}$ can be viewed as a collection of homogeneous spaces G/H of varying dimension: some finite-dimensional, others infinite-dimensional. However, except in the case of smooth measures discussed in [18], these H are not known to be "infinite-dimensional Lie groups" and, to our knowledge, the corresponding homogeneous spaces are not known to be "infinite-dimensional manifolds" in the sense of the above references.

4. TANGENT AND COTANGENT BUNDLES

We now define some further elements of calculus on $\mathcal{M}$. As opposed to Section 3, the definitions and statements made here are completely rigorous. We will often refer back to the ideas of Section 3 and to the appendix, however, to explain the geometric intuition underlying this theory.

4.1. **Push-forward operations on $\mathcal{M}$ and $T\mathcal{M}$.** The following results concern the push-forward operation on $\mathcal{M}$.

Lemma 4.1. *If $\phi : \mathbb{R}^D \to \mathbb{R}^D$ is a Lipschitz map with Lipschitz constant $Lip\,\phi$ then $\phi_\# : \mathcal{M} \to \mathcal{M}$ is also a Lipschitz map with the same Lipschitz constant.*

Proof: Let $\mu, \nu \in \mathcal{M}$. Note that if $u(x) + v(y) \leq |x - y|^2$ for all $x, y \in \mathbb{R}^D$ then

$$u \circ \phi(a) + v \circ \phi(b) \leq |\phi(a) - \phi(b)|^2 \leq (Lip\,\phi)^2 |a - b|^2.$$

This, together with Equation 2.2, yields

$$(4.1) \qquad \int_{\mathbb{R}^D} u\, d\phi_{\#}\mu + \int_{\mathbb{R}^D} v\, d\phi_{\#}\nu = \int_{\mathbb{R}^D} u \circ \phi\, d\mu + \int_{\mathbb{R}^D} v \circ \phi\, d\nu \leq (Lip\,\phi)^2 W_2^2(\mu, \nu).$$

We maximize the expression on the left hand side of Equation 4.1 over the set of pairs (u, v) such that $u(x) + v(y) \leq |x - y|^2$ for all $x, y \in \mathbb{R}^D$. Then we use again Equation 2.2 to conclude the proof. $\hspace{2cm}$ QED.

Lemma 4.2. *For any $\mu \in \mathcal{M}$ and $\phi \in Diff_c(\mathbb{R}^D)$, the map $\phi_* : \mathcal{X}_c \to \mathcal{X}_c$ has a unique continuous extension $\phi_* : L^2(\mu) \to L^2(\phi_{\#}\mu)$. Furthermore $\phi_*\big(Ker(div_\mu)\big) \subseteq Ker(div_{\varphi_{\#}\mu})$. Thus ϕ_* induces a continuous map $\phi_* : T_\mu\mathcal{M} \to T_{\phi_{\#}\mu}\mathcal{M}$.*

Proof: Let $\mu \in \mathcal{M}$, $\phi \in \text{Diff}_c(\mathbb{R}^D)$, $f \in C_c^\infty(\mathbb{R}^D)$ and let $X \in \text{Ker}(div_\mu)$. If C_ϕ is the L^∞-norm of $\nabla\phi$ we have $||\phi_* X||_{\phi_{\#}\mu} \leq C_\phi ||X||_\mu$. Hence ϕ_* admits a unique continuous linear extension. Furthermore

$$\int_{\mathbb{R}^D} df(\varphi_* X)\, d\varphi_{\#}\mu = \int_{\mathbb{R}^D} df_{|\varphi}(\varphi_* X_{|\varphi})\, d\mu = \int_{\mathbb{R}^D} df_{|\varphi}(\nabla\varphi \cdot X)\, d\mu = \int_{\mathbb{R}^D} d(f \circ \varphi)(X)\, d\mu = 0.$$

$$\text{QED.}$$

Lemma 4.3. *Let $\sigma \in AC_2(a, b; \mathcal{M})$ and let v be a velocity for σ. Let $\varphi \in Diff_c(\mathbb{R}^D)$. Then $t \to \varphi_{\#}(\sigma_t) \in AC_2(a, b; \mathcal{M})$ and $\varphi_* v$ is a velocity for $\varphi_{\#}\sigma$.*

Proof: If $a < s < t < b$ then, by Lemma 4.1, $W_2(\varphi_{\#}\sigma_t, \varphi_{\#}\sigma_s) \leq (Lip\,\varphi)\, W_2(\sigma_t, \sigma_s)$. Since $\sigma \in AC_2(a, b; \mathcal{M})$ one concludes that $\varphi_{\#}(\sigma) \in AC_2(a, b; \mathcal{M})$. If $f \in C_c^\infty((a, b) \times \mathbb{R}^D)$ we have

$$\int_a^b \int_{\mathbb{R}^D} \Big(\frac{\partial f_t}{\partial t} + df_t(\phi_* v_t) \Big) d(\varphi_{\#}\sigma_t) dt = \int_a^b \int_{\mathbb{R}^D} \Big(\frac{\partial f_t}{\partial t} \circ \varphi + (df_t(\phi_* v_t)) \circ \varphi \Big) d\sigma_t dt$$

$$= \int_a^b \int_{\mathbb{R}^D} \Big(\frac{\partial (f_t \circ \varphi)}{\partial t} + d(f_t \circ \varphi)(v_t) \Big) d\sigma_t dt$$

$$= 0.$$

To obtain the last equality we have used that $(t, x) \to f(t, \varphi(x))$ is in $C_c^\infty((a, b) \times \mathbb{R}^D)$.

$$\text{QED.}$$

4.2. **Differential forms on $\mathcal{M}$.** Recall from Definition 2.5 that the tangent bundle $T\mathcal{M}$ of $\mathcal{M}$ is defined as the union of all spaces $T_\mu\mathcal{M}$, for $\mu \in \mathcal{M}$. We now define the *pseudo tangent bundle* $\mathcal{T}\mathcal{M}$ to be the union of all spaces $L^2(\mu)$. Analogously, the union of the dual spaces $T_\mu^*\mathcal{M}$ defines the *cotangent bundle* $T^*\mathcal{M}$; we define the *pseudo cotangent bundle* $\mathcal{T}^*\mathcal{M}$ to be the union of the dual spaces $L^2(\mu)^*$.

It is clear from the definitions that we can think of $T\mathcal{M}$ as a subbundle of $\mathcal{T}\mathcal{M}$. Decomposition 2.5 allows us also to define an injection $T^*\mathcal{M} \to \mathcal{T}^*\mathcal{M}$ by extending any covector $T_\mu\mathcal{M} \to \mathbb{R}$ to be zero on the complement of $T_\mu\mathcal{M}$ in $L^2(\mu)$. In this sense we can also think of $T^*\mathcal{M}$ as a subbundle of $\mathcal{T}^*\mathcal{M}$. The projections π_μ from Section 2.3 combine to define a surjection $\pi : \mathcal{T}\mathcal{M} \to T\mathcal{M}$. Likewise, restriction yields a surjection $\mathcal{T}^*\mathcal{M} \to T^*\mathcal{M}$.

Remark 4.4. The above constructions make heavy use of the Hilbert structure on $L^2(\mu)$. Following the point of view of Remark 2.7 and Section 3.2, *i.e.* emphasizing the differential, rather than the Riemannian, structure of $\mathcal{M}$ one could decide to define $T_\mu\mathcal{M}$ as $L^2(\mu)/\mathrm{Ker}(div_\mu)$. Then the projections $\pi_\mu : L^2(\mu) \to T_\mu\mathcal{M}$ would still define by duality an injection $T^*\mathcal{M} \to \mathcal{T}^*\mathcal{M}$: this would identify $T^*\mathcal{M}$ with the *annihilator* of $\mathrm{Ker}(div_\mu)$ in $L^2(\mu)$. However there would be no natural injection $T\mathcal{M} \to \mathcal{T}\mathcal{M}$ nor any natural surjection $\mathcal{T}^*\mathcal{M} \to T^*\mathcal{M}$.

Definition 4.5. A *1-form* on $\mathcal{M}$ is a section of the cotangent bundle $T^*\mathcal{M}$, *i.e.* a collection of maps $\mu \mapsto \Lambda_\mu \in T_\mu^*\mathcal{M}$. A *pseudo 1-form* is a section of the pseudo cotangent bundle $\mathcal{T}^*\mathcal{M}$. Analogously, a *2-form* on $\mathcal{M}$ is a collection of alternating multilinear maps

$$\mu \mapsto \Lambda_\mu : T_\mu\mathcal{M} \times T_\mu\mathcal{M} \to \mathbb{R}.$$

A *pseudo 2-form* is a collection of alternating multilinear maps

$$\mu \mapsto \bar{\Lambda}_\mu : L^2(\mu) \times L^2(\mu) \to \mathbb{R}.$$

It is natural (but in practice sometimes too strong) to further assume that each Λ_μ (or $\bar{\Lambda}_\mu$) satisfies a continuity assumption such as the following: there exists $c_\mu \in \mathbb{R}$ such that

$$|\Lambda_\mu(X_1, X_2)| \le c_\mu \|X_1\|_\mu \cdot \|X_2\|_\mu.$$

For $k = 1, 2$ we let $\Lambda^k\mathcal{M}$ (respectively, $\bar{\Lambda}^k\mathcal{M}$) denote the space of k-forms (respectively, pseudo k-forms). We define a *0-form* to be a function $F : \mathcal{M} \to \mathbb{R}$.

Notice that, for $k = 1$, continuity implies that any 1-form is uniquely defined by its values on any dense subset of $T_\mu\mathcal{M}$, *e.g.* on the dense subset defined by smooth gradient vector fields. The analogue holds also for pseudo forms and for $k = 2$, as long as the continuity condition holds. As above, extension defines a natural injection

(4.2) $$\Lambda^k\mathcal{M} \to \bar{\Lambda}^k\mathcal{M}, \quad \Lambda \mapsto \bar{\Lambda},$$

i.e. to every k-form one can associate a canonical pseudo k-form. Conversely, restriction defines a surjection $\bar{\Lambda}^k\mathcal{M} \to \Lambda^k\mathcal{M}$.

Since $T_\mu\mathcal{M}$ is a Hilbert space, by the Riesz representation theorem every 1-form Λ_μ on $T_\mu\mathcal{M}$ can be written $\Lambda_\mu(Y) = \int_{\mathbb{R}^D}\langle A_\mu, Y\rangle d\mu$ for a unique $A_\mu \in T_\mu\mathcal{M}$ and all $Y \in T_\mu\mathcal{M}$. The analogous fact is true also for pseudo 1-forms.

Example 4.6. Any $f \in C_c^\infty$ defines a function on $\mathcal{M}$, *i.e.* a 0-form, as follows:

$$F(\mu) := \int_{\mathbb{R}^D} f d\mu.$$

We will refer to these as the *linear* functions on $\mathcal{M}$, in that the natural extension to the space $(C_c^\infty)^*$ defines a function which is linear with respect to μ.

Any $\bar{A} \in \mathcal{X}_c$ defines a pseudo 1-form on $\mathcal{M}$ as follows:

(4.3) $$\bar{\Lambda}_\mu(X) := \int_{\mathbb{R}^D}\langle \bar{A}, X\rangle d\mu.$$

We will refer to these as the *linear* pseudo 1-forms. Notice that if $\bar{A} = \nabla f$ for some $f \in C_c^\infty$ then $\bar{\Lambda}$ is actually a 1-form.

Any bounded field $B = B(x)$ on $\mathbb{R}^D$ of $D \times D$ matrices defines a *linear* pseudo 2-form via

$$(4.4) \qquad \hat{B}_\mu(X, Y) := \int_{\mathbb{R}^D} \langle BX, Y \rangle d\mu.$$

Remark 4.7. When $k = 1$ Hölder's inequality for the product of two functions shows that any vector field $\bar{A}_\mu \in L^p(\mu)$, for $p \in [2, \infty]$, defines a continuous map $L^2(\mu) \to \mathbb{R}$ as in Equation 4.3. When $k = 2$ Hölder's inequality for the product of three functions shows that any field of matrices $B_\mu \in L^\infty(\mu)$ defines a continuous map $L^2(\mu) \times L^2(\mu) \to \mathbb{R}$ as in Equation 4.4. However, for $k \geq 3$ there do not exist analogous constructions of (non-trivial) continuous k-linear maps $L^2(\mu) \times \cdots \times L^2(\mu) \to \mathbb{R}$. It is for this reason that we restrict our attention to the case $k \leq 2$. In any case, this is sufficient for the applications of interest to us.

As in Section A.2, the action of $\mathrm{Diff}_c(\mathbb{R}^D)$ on $\mathcal{M}$ can be lifted to forms and pseudo forms as follows.

Definition 4.8. For $k = 1, 2$, let $\bar{\Lambda}$ be a pseudo k-form on $\mathcal{M}$. Then any $\phi \in \mathrm{Diff}_c(\mathbb{R}^D)$ defines a *pull-back* k-multilinear map $\phi^* \bar{\Lambda}$ on $\mathcal{M}$ as follows:

$$(\phi^* \bar{\Lambda})_\mu(X_1, \ldots, X_k) := \bar{\Lambda}_{\phi_\# \mu}(\phi_* X_1, \ldots, \phi_* X_k).$$

It follows from Lemma 4.2 that the push-forward operation preserves Decomposition 2.5. This implies that the pull-back preserves the space of k-forms, *i.e.* the pull-back of a k-form is a k-form.

Definition 4.9. Let $F : \mathcal{M} \to \mathbb{R}$ be a function on $\mathcal{M}$. We say that $\xi \in L^2(\mu)$ belongs to the *subdifferential* $\partial_- F(\mu)$ if

$$F(\nu) \geq F(\mu) + \sup_{\gamma \in \Gamma_o(\mu, \nu)} \iint_{\mathbb{R}^D \times \mathbb{R}^D} \langle \xi(x), y - x \rangle \, d\gamma(x, y) + o(W_2(\mu, \nu)),$$

as $\nu \to \mu$. If $-\xi \in \partial_-(-F)(\mu)$ we say that ξ belongs to the *superdifferential* $\partial^+ F(\mu)$.

If $\xi \in \partial_- F(\mu) \cap \partial^+ F(\mu)$ then, for any $\gamma \in \Gamma_o(\mu, \nu)$,

$$(4.5) \qquad F(\nu) = F(\mu) + \iint_{\mathbb{R}^D \times \mathbb{R}^D} \langle \xi(x), y - x \rangle \, d\gamma(x, y) + o(W_2(\mu, \nu)).$$

If such ξ exists we say that F is *differentiable* at μ and we define the *gradient vector* $\nabla_\mu F := \pi_\mu(\xi)$. Using barycentric projections (cf. [5] Definition 5.4.2) one can show that, for $\gamma \in \Gamma_o(\mu, \nu)$,

$$\iint_{\mathbb{R}^D \times \mathbb{R}^D} \langle \xi(x), y - x \rangle \, d\gamma(x, y) = \iint_{\mathbb{R}^D \times \mathbb{R}^D} \langle \pi_\mu(\xi)(x), y - x \rangle \, d\gamma(x, y).$$

Thus $\pi_\mu(\xi) \in \partial_- F(\mu) \cap \partial^+ F(\mu) \cap T_\mu \mathcal{M}$ and it satisfies the analogue of Equation 4.5. It can be shown that the gradient vector is unique, *i.e.* that $\partial_- F(\mu) \cap \partial^+ F(\mu) \cap T_\mu \mathcal{M} = \{\pi_\mu(\xi)\}$.

Finally, if the gradient vector exists for every $\mu \in \mathcal{M}$ we can define the *differential* or *exterior derivative* of F to be the 1-form dF determined, for any $\mu \in \mathcal{M}$ and $Y \in T_\mu \mathcal{M}$, by $dF(\mu)(Y) := \int_{\mathbb{R}^D} \langle \nabla_\mu F, Y \rangle \, d\mu$. To simplify the notation we will sometimes write $Y(F)$ rather then $dF(Y)$.

Remark 4.10. Assume $F : \mathcal{M} \to \mathbb{R}$ is differentiable. Given $X \in \nabla C_c^\infty(\mathbb{R}^D)$, let ϕ_t denote the flow of X. Fix $\mu \in \mathcal{M}$.

(i) Set $\nu_t := (Id + tX)_{\#}\mu$. Then

$$F(\nu_t) = F(\mu) + t \int_{\mathbb{R}^D} \langle \nabla_\mu F, X \rangle d\mu + o(t).$$

(ii) Set $\mu_t := \phi_{t\#}\mu$. If $||\nabla_\mu F(\mu)||_\mu$ is bounded on compact subsets of $\mathcal{M}$ then

$$F(\mu_t) = F(\mu) + t \int_{\mathbb{R}^D} \langle \nabla_\mu F, X \rangle d\mu + o(t).$$

Proof: The proof of (i) is a direct consequence of Equation 4.5 and of the fact that, if $r > 0$ is small enough, $\left(Id \times (Id + tX)\right)_{\#}\mu \in \Gamma_o(\mu, \nu_t)$ for $t \in [-r, r]$.

To prove (ii), set
$$A(s,t) := (1-s)(Id + tX) + s\phi_t.$$
Notice that $||\phi_t - Id - tX||_\mu \leq t^2||(\nabla X)X||_\infty$ and that $(s,t) \to m(s,t) := A(s,t)_{\#}\mu$ defines a continuous map of the compact set $[0,1] \times [-r, r]$ into $\mathcal{M}$. Hence the range of m is compact so $||\nabla_\mu F(\mu)||_\mu$ is bounded there by a constant C. One can use elementary arguments to conclude that F is C-Lipschitz on the range of m, cf. [25] for details. Let $\gamma_t := \left((Id + tX) \times \phi_t\right)_{\#}\mu$. We have $\gamma_t \in \Gamma(\nu_t, \mu_t)$ so $W_2(\mu_t, \nu_t) \leq ||\phi_t - Id - tX||_\mu = O(t^2)$. We conclude that

$$|F(\nu_t) - F(\mu_t)| \leq CW_2(\mu_t, \nu_t) = O(t^2).$$

This, together with (i), yields (ii).   QED.

Example 4.11. Fix $f \in C_c^\infty$ and let $F : \mathcal{M} \to \mathbb{R}$ be the corresponding linear function, as in Example 4.6. Then F is differentiable with gradient $\nabla_\mu F \equiv \nabla f$. Thus dF is a linear 1-form on $\mathcal{M}$. Viceversa, according to our definitions every linear 1-form Λ is *exact*. In other words, if $\Lambda_\mu(X) = \int_{\mathbb{R}^D} \langle A, X \rangle d\mu$ for some $A = \nabla f$ then $\Lambda = dF$ for $F(\mu) := \int_{\mathbb{R}^D} f \, d\mu$.

Definition 4.12. Let $\bar{\Lambda}$ be a pseudo 1-form on $\mathcal{M}$. We say that $\bar{\Lambda}$ is *differentiable* if the following two conditions hold:

(i) For all $X \in \mathcal{X}_c$, the function $\bar{\Lambda}(X) : \mathcal{M} \to \mathbb{R}$ is differentiable. We can then define the *exterior derivative* of $\bar{\Lambda}$ on pairs $X, Y \in \mathcal{X}_c$ by setting

(4.6) $$d\bar{\Lambda}(X, Y) := X\bar{\Lambda}(Y) - Y\bar{\Lambda}(X) - \bar{\Lambda}([X, Y]).$$

(ii) For all $\mu \in \mathcal{M}$, $d\bar{\Lambda}_\mu$ is continuous when restricted to ∇C_c^∞, *i.e.* there exists $c_\mu \in \mathbb{R}$ such that

$$|d\bar{\Lambda}_\mu(\nabla f, \nabla g)| \leq c_\mu ||\nabla f||_\mu \cdot ||\nabla g||_\mu, \quad \text{for all } \nabla f, \nabla g \in \nabla C_c^\infty.$$

Notice that condition (ii) implies that $d\bar{\Lambda}_\mu$ has a unique extension to $T_\mu\mathcal{M} \times T_\mu\mathcal{M}$.

Let Λ be a 1-form on $\mathcal{M}$. Let $\bar{\Lambda}$ denote the associated pseudo 1-form, as in Equation 4.2. We say that Λ is *differentiable* if $\bar{\Lambda}$ is differentiable. We can then define its *exterior derivative* by setting $d\Lambda := d\bar{\Lambda}$.

Remark 4.13. The assumption that $d\bar{\Lambda}$ satisfies the continuity assumption (ii) on ∇C_c^∞ implies that some form of cancelling occurs to eliminate first-order terms as in Equation A.11, cf. Remark A.7. Notice that $d\bar{\Lambda}$, restricted to $T_\mu\mathcal{M} \times T_\mu\mathcal{M}$, is a well-defined 2-form. On the other hand, Example 4.15 shows that it is not natural to impose a continuity assumption on $\mathcal{X}_c$ so $d\bar{\Lambda}$ does not in general extend to a uniquely defined pseudo 2-form.

If Λ is differentiable in the above sense, it is natural to ask if $d\bar{\Lambda}_\mu(X, \cdot) = 0$ for any $X \in \text{Ker}(div_\mu)$. It is not clear that this is the case.

Remark 4.14. One could also define a notion of differentiability for 1-forms by testing Λ only against gradient vector fields $\nabla f \in \nabla C_c^\infty$. This is clearly a weaker condition, which would yield a very poor understanding of the differentiability of the associated pseudo 1-form $\bar{\Lambda}$. Indeed, assume Λ is differentiable in the weaker sense and choose $X \in \mathcal{X}_c$. Then $\bar{\Lambda}(X) = \Lambda(\pi_\mu(X))$ and $\pi_\mu(X)$ depends on μ. In particular, the differentiability of $\bar{\Lambda}$ is now related to the smoothness of the projection operators $\mu \mapsto \pi_\mu$. We will avoid this notion, using instead the stronger definition given in Definition 4.12.

Example 4.15. Assume $\bar{\Lambda}$ is a linear pseudo 1-form, *i.e.* $\bar{\Lambda}(\cdot) = \int_{\mathbb{R}^D} \langle \bar{A}, \cdot \rangle d\mu$ for some $\bar{A} \in \mathcal{X}_c$. Then $\bar{\Lambda}$ is differentiable and, $\forall X, Y \in \mathcal{X}_c$,

$$d\bar{\Lambda}(X,Y) = \int_{\mathbb{R}^D} < \pi_\mu(\nabla \bar{A}^T \cdot Y + \nabla Y^T \cdot \bar{A}), X > - < \pi_\mu(\nabla \bar{A}^T \cdot X + \nabla X^T \cdot \bar{A}), Y > d\mu$$
$$- \int_{\mathbb{R}^D} < \nabla Y \cdot X - \nabla X \cdot Y, \bar{A} > d\mu.$$

If $X, Y \in T_\mu \mathcal{M}$ then $d\bar{\Lambda}(X,Y) = \int_{\mathbb{R}^D} < (\nabla \bar{A} - \nabla \bar{A}^T)X, Y > d\mu$.

Proof: Define $F_X : \mathcal{M} \to \mathbb{R}$ by

$$F_X(\mu) := \bar{\Lambda}(X) = \int_{\mathbb{R}^D} < \bar{A}, X > d\mu.$$

Let $\mu, \nu \in \mathcal{M}$ and $\gamma \in \Gamma_o(\mu, \nu)$. Then

$$F_X(\nu) - F_X(\mu) = \int_{\mathbb{R}^D} < \bar{A}, X > d\nu - \int_{\mathbb{R}^D} < \bar{A}, X > d\mu$$
$$= \iint_{\mathbb{R}^D \times \mathbb{R}^D} < \bar{A}(y), X(y) > - < \bar{A}(x), X(x) > d\gamma(x, y).$$

Set $\phi := < \bar{A}, X >$. Then $\phi \in C_c^\infty(\mathbb{R}^D)$ so

$$(4.7) \qquad F_X(\nu) - F_X(\mu) = \iint_{\mathbb{R}^D \times \mathbb{R}^D} \phi(y) - \phi(x) d\gamma(x, y)$$
$$= \iint_{\mathbb{R}^D \times \mathbb{R}^D} (< \nabla \phi(x), y - x > + O(|x - y|^2) d\gamma(x, y)$$
$$= \iint_{\mathbb{R}^D \times \mathbb{R}^D} < \nabla \phi(x), y - x > d\gamma(x, y) + o(W_2(\mu, \nu)).$$

Equation 4.7 shows that $F_X : \mathcal{M} \to \mathbb{R}$ is differentiable and that $\nabla_\mu F_X = \pi_\mu(\nabla \phi)$. Thus

$$(4.8) \qquad Y \bar{\Lambda}(X) := dF_X(Y) = \int_{\mathbb{R}^D} < \nabla_\mu F_X, Y > d\mu$$
$$= \int_{\mathbb{R}^D} < \pi_\mu(\nabla \bar{A}^T \cdot X + \nabla X^T \cdot \bar{A}), Y > d\mu.$$

Analogously,

$$(4.9) \qquad X\bar{\Lambda}(Y) := dF_Y(X) = \int_{\mathbb{R}^D} <\nabla_\mu F_Y, X> d\mu$$

$$= \int_{\mathbb{R}^D} <\pi_\mu(\nabla \bar{A}^T \cdot Y + \nabla Y^T \cdot \bar{A}), X> d\mu.$$

We combine Equations 4.8 and 4.9 to get

$$d\bar{\Lambda}(X,Y) := X\bar{\Lambda}(Y) - Y\bar{\Lambda}(X) - \bar{\Lambda}([X,Y])$$

$$(4.10) \qquad = \int_{\mathbb{R}^D} <\pi_\mu(\nabla \bar{A}^T \cdot Y + \nabla Y^T \cdot \bar{A}), X> d\mu$$

$$- \int_{\mathbb{R}^D} <\pi_\mu(\nabla \bar{A}^T \cdot X + \nabla X^T \cdot \bar{A}), Y> d\mu$$

$$- \int_{\mathbb{R}^D} <\nabla Y \cdot X - \nabla X \cdot Y, \bar{A}> d\mu.$$

If $X, Y \in T_\mu \mathcal{M}$ then Equation 4.10 simplifies to

$$d\bar{\Lambda}(X,Y) = \int_{\mathbb{R}^D} <(\nabla \bar{A}^T \cdot Y + \nabla Y^T \cdot \bar{A}), X> - <(\nabla \bar{A}^T \cdot X + \nabla X^T \cdot \bar{A}), Y> d\mu$$

$$- \int_{\mathbb{R}^D} <\nabla Y \cdot X - \nabla X \cdot Y, \bar{A}> d\mu$$

$$= \int_{\mathbb{R}^D} <(\nabla \bar{A} - \nabla \bar{A}^T)X, Y> d\mu.$$

$$\text{QED.}$$

In Lemma 5.12 we will generalize this result to the class of *regular* pseudo 1-forms.

4.3. **Discussion.** As explained in Section 3.2, we can think of $\mathcal{M}$ as the union of smooth manifolds $\mathcal{O}$. Each tangent space $T_\mu \mathcal{M}$ should then be thought of as the tangent space of $\mathcal{O}$ at the point μ. Our notion of k-form Λ is defined in terms of the dual tangent spaces, so each $\Lambda_{|\mathcal{O}}$ is, at least formally, a k-form on a smooth manifold in the usual sense. The logic behind our definition of the operator d on 1-forms is as follows. As seen in Section 3.2, any $X \in \mathcal{X}_c$ defines a fundamental vector field on $\mathcal{O}$ (or on $\mathcal{M}$). In particular we can think of the construction of fundamental vector fields as a canonical way of extending given tangent vectors X, Y at any point $\mu \in \mathcal{O}$ to global tangent vector fields on $\mathcal{O}$. Combining Remark A.13 with Lemma A.19 shows that the construction of fundamental vector fields determines a Lie algebra homomorphism $\mathcal{X}_c \to \mathcal{X}(\mathcal{O})$. Equation 4.6 thus mimics Equation A.11 for $k = 1$. In Section 5.4 we will study the corresponding first cohomology group. We can think of this as the *de Rham cohomology* of the manifold $\mathcal{O}$.

The notion of pseudo k-form is less standard, but also very natural. The finite-dimensional analogue of this notion is explained in Section A.3. Roughly speaking, *i.e.* up to L^2_μ-closure, if we restrict our space of pseudo k-forms to any manifold $\mathcal{O}$ we obtain the space of maps $\mathcal{O} \to \Lambda^k(\mathfrak{g})$, where $\mathfrak{g} = \mathcal{X}_c$. Our definition of the operator d, given in Equation 4.6, should now be compared to Equation A.19. Notice that the sign discrepancy between these equations is explained by the fact that the Lie bracket on $\mathfrak{g}$ is the opposite of the usual Lie bracket on $\mathcal{X}_c$, cf. Lemma A.19. In this setting the key point is that each manifold $\mathcal{O}$ is

actually the orbit of a group action. More specifically, we can identify it with a quotient of the group $\text{Diff}_c(\mathbb{R}^D)$. Proposition A.17 then shows that pseudo k-forms are actually k-forms on the group, rather than on the manifold, endowed with a special "invariance" property. In some sense the corresponding cohomology is more closely related to the orbit structure of the manifold $\mathcal{O}$ than to its topological structure. However Proposition A.17 shows that, at least in finite dimensions, there is a simple relation between this *invariant cohomology* and the usual de Rham cohomology of the manifold: for $k = 1$, the latter is a subgroup of the former. Proposition A.17 also shows that the operation of Equation 4.2 is very natural from this point of view: up to the appropriate identifications, it coincides with the standard pull-back operation from k-forms on the quotient of the group to k-forms on the group.

It may be useful to emphasize that the identification between the orbit $\mathcal{O}$ and the quotient space is not canonical. The details involved in changing this identification are explained in Section A.2.

5. Calculus of pseudo differential 1-forms

Given a 1-form α on a finite-dimensional manifold, Green's formula relates the integral of $d\alpha$ along a surface to the integral of α along the boundary curves. In Section 5.1 we show that an analogous result for $\mathcal{M}$ is rather simple if both the form and the surface satisfy certain regularity conditions. The conditions we need to impose on the form are rather mild: we investigate these in Sections 5.2 and 5.3, developing a general theory of *regular* pseudo 1-forms. The conditions on the surface, instead, are very strong. In Section 5.4 we thus prove a second version of Green's formula, valid only for certain surfaces we call *annuli*. For these surfaces we need no extra regularity conditions, and Green's formula then leads to a proof that every closed regular 1-form is exact.

5.1. **Green's formula for smooth surfaces and 1-forms.** Let Λ be a differentiable 1-form on $\mathcal{M}$ in the sense of Definition 4.12. Let $\bar{\Lambda}$ denote the associated pseudo 1-form in the sense of Equation 4.2. Set $||\Lambda_\mu|| := \sup_v \{\Lambda_\mu(v) : v \in T_\mu \mathcal{M}, ||v||_\mu \leq 1\}$. We assume that, for all compact subsets $\mathcal{K} \subset \mathcal{M}$,

$$(5.1) \qquad \qquad \sup_{\mu \in \mathcal{K}} ||\Lambda_\mu|| < \infty.$$

We also assume that for all compact subsets $\mathcal{K} \subset \mathcal{M}$ there exists a constant $C_\mathcal{K}$ such that, for all $\mu, \nu \in \mathcal{K}$ and $u \in C_b(\mathbb{R}^D, \mathbb{R}^D)$ such that ∇u is bounded,

$$(5.2) \qquad \qquad |\bar{\Lambda}_\nu(u) - \bar{\Lambda}_\mu(u)| \leq C_\mathcal{K} W_2(\mu, \nu)(||u||_\infty + ||\nabla u||_\infty).$$

Set $||d\Lambda_\mu||$ to be the smallest nonnegative number c_μ such that

$$|d\Lambda_\mu(\nabla f, \nabla g)| \leq c_\mu ||\nabla f||_\mu \cdot ||\nabla g||_\mu, \quad \text{for all } \nabla f, \nabla g \in \nabla C_c^\infty.$$

Now let $S : [0,1] \times [0,T] \to \mathcal{M}$ denote a map satisfying the following three regularity conditions:

(i) For each $s \in [0,1]$, $S(s, \cdot) \in AC_2(0, T; \mathcal{M})$ and, for each $t \in [0, T]$, $S(\cdot, t) \in AC_2(0, 1; \mathcal{M})$.

(ii) Let $v(s, \cdot, \cdot)$ denote the velocity of minimal norm for $S(s, \cdot)$ and $w(\cdot, t, \cdot)$ denote the velocity of minimal norm for $S(\cdot, t)$. We assume that $v, w \in C^2([0,1] \times [0,T] \times \mathbb{R}^D, \mathbb{R}^D)$ and that their derivatives up to third order are bounded. We further assume that v and w are gradient vector fields so that $\partial_s v$ and $\partial_t w$ are also gradients: this implies that Λ and $\bar{\Lambda}$ coincide when evaluated on these fields.

(iii) S takes values in the set of absolutely continuous measures. More specifically, $S(s,t) = \rho(s,t,\cdot)\mathcal{L}^D$ for some $\rho \in C^1([0,1] \times [0,T] \times \mathbb{R}^D)$ which is bounded with bounded derivatives.

Using Remark 2.11, Proposition 2.12 and the bound on v, w and on their derivatives, we find that S is 1/2-Hölder continuous. Hence its range is compact so $||\Lambda_{S(s,t)}||$ is bounded. We then use Equations 5.1, 5.2 and Taylor expansions for w_{t+h}^s and v_t^{s+h} to obtain that

$$(5.3) \qquad \partial_t\Big(\Lambda_{S(s,t)}(w_t^s)\Big)_{|\bar{s},\bar{t}} = v_{\bar{t}}^{\bar{s}}(\Lambda_{S(s,t)}(w_{\bar{t}}^{\bar{s}})) + \Lambda_{S(\bar{s},\bar{t})}(\partial_t w_t^s),$$

where we use the notation of Definition 4.9. Analogously,

$$(5.4) \qquad \partial_s\Big(\Lambda_{S(s,t)}(v_t^s)\Big)_{|\bar{s},\bar{t}} = w_{\bar{t}}^{\bar{s}}(\Lambda_{S(s,t)}(v_{\bar{t}}^{\bar{s}})) + \Lambda_{S(\bar{s},\bar{t})}(\partial_s v_t^s).$$

Lemma 5.1. *For $(s,t) \in (0,1) \times (0,T)$ we have $(\partial_t w_t^s - \partial_s v_t^s) - [w_t^s, v_t^s] \in Ker(div_{S(s,t)})$.*

Proof: We have, in the sense of distributions,

$$(5.5) \qquad \partial_t \rho_t^s + \nabla \cdot (\rho_t^s v_t^s) = 0, \quad \partial_s \rho_t^s + \nabla \cdot (\rho_t^s w_t^s) = 0$$

and so

$$\nabla \cdot \partial_s(\rho_t^s v_t^s) = -\partial_s \partial_t \rho_t^s = \nabla \cdot (\partial_t \rho_t^s w_t^s).$$

We use that ρ, v and w are smooth to conclude that

$$\nabla \cdot \Big(v_t^s \partial_s \rho_t^s + \rho_t^s \partial_s v_t^s\Big) = \nabla \cdot \Big(w_t^s \partial_t \rho_t^s + \rho_t^s \partial_t w_t^s\Big).$$

This implies that if $\varphi \in C_c^\infty(\mathbb{R}^D)$ then

$$(5.6) \qquad \int_{\mathbb{R}^D} \langle \nabla\varphi, v_t^s \partial_s \rho_t^s + \rho_t^s \partial_s v_t^s \rangle = \int_{\mathbb{R}^D} \langle \nabla\varphi, w_t^s \partial_t \rho_t^s + \rho_t^s \partial_t w_t^s \rangle.$$

We use again that ρ, v and w are smooth to obtain that Equation 5.5 holds pointwise. Hence, Equation 5.6 implies

$$\int_{\mathbb{R}^D} \langle \nabla\varphi, -v_t^s \nabla \cdot (\rho_t^s w_t^s) + \rho_t^s \partial_s v_t^s \rangle = \int_{\mathbb{R}^D} \langle \nabla\varphi, -w_t^s \nabla \cdot (\rho_t^s v_t^s) + \rho_t^s \partial_t w_t^s \rangle.$$

Rearranging, this leads to

$$\int_{\mathbb{R}^D} \langle \nabla\varphi, \partial_s v_t^s - \partial_t w_t^s \rangle \rho_t^s d\mathcal{L}^D = \int_{\mathbb{R}^D} \langle \nabla\varphi, v_t^s \rangle \nabla \cdot (\rho_t^s w_t^s) - \langle \nabla\varphi, w_t^s \rangle \nabla \cdot (\rho_t^s v_t^s).$$

Integrating by parts and substituting $\rho_t^s \mathcal{L}^D$ with $S(s,t)$ we obtain

$$\int_{\mathbb{R}^D} \langle \nabla\varphi, \partial_s v_t^s - \partial_t w_t^s \rangle dS(s,t)$$

$$= \int_{\mathbb{R}^D} \Big(\langle \nabla^2 \varphi w_t^s + (\nabla w_t^s)^T \nabla\varphi, v_t^s \rangle - \langle \nabla^2 \varphi v_t^s + (\nabla v_t^s)^T \nabla\varphi, w_t^s \rangle\Big) dS(s,t)$$

$$= \int_{\mathbb{R}^D} \langle \nabla\varphi, [v_t^s, w_t^s] \rangle dS(s,t).$$

Since $\varphi \in C_c^\infty(\mathbb{R}^D)$ is arbitrary, the proof is finished.   QED.

Proposition 5.2. *For each $t \in (0, T)$ and $s \in (0, 1)$ we have*

$$\partial_t \Big(\Lambda_{S(s,t)}(w_t^s) \Big) - \partial_s \Big(\Lambda_{S(s,t)}(v_t^s) \Big) = d\Lambda_{S(s,t)}(v_t^s, w_t^s).$$

Proof: We use Definition 4.12 and Equations 5.3, 5.4 to find

$$
\begin{aligned}
d\Lambda_{S(\bar{s},\bar{t})}(v_{\bar{t}}^{\bar{s}}, w_{\bar{t}}^{\bar{s}}) &= d\bar{\Lambda}_{S(\bar{s},\bar{t})}(v_{\bar{t}}^{\bar{s}}, w_{\bar{t}}^{\bar{s}}) \\
&= v_{\bar{t}}^{\bar{s}}(\bar{\Lambda}_{S(s,t)}(w_{\bar{t}}^{\bar{s}})) - w_{\bar{t}}^{\bar{s}}(\bar{\Lambda}_{S(s,t)}(v_{\bar{t}}^{\bar{s}})) - \bar{\Lambda}_{S(\bar{s},\bar{t})}([v_{\bar{t}}^{\bar{s}}, w_{\bar{t}}^{\bar{s}}]) \\
&= v_{\bar{t}}^{\bar{s}}(\Lambda_{S(s,t)}(w_{\bar{t}}^{\bar{s}})) - w_{\bar{t}}^{\bar{s}}(\Lambda_{S(s,t)}(v_{\bar{t}}^{\bar{s}})) - \bar{\Lambda}_{S(\bar{s},\bar{t})}([v_t^s, w_t^s]) \\
&= \partial_t \Big(\Lambda_{S(s,t)}(w_t^s) \Big)_{|\bar{s},\bar{t}} - \Lambda_{S(\bar{s},\bar{t})}(\partial_t w_t^s) - \partial_s \Big(\Lambda_{S(s,t)}(v_t^s) \Big)_{|\bar{s},\bar{t}} \\
&\quad + \Lambda_{S(\bar{s},\bar{t})}(\partial_s v_t^s) - \bar{\Lambda}_{S(\bar{s},\bar{t})}([v_t^s, w_t^s]) \\
&= \partial_t \Big(\Lambda_{S(s,t)}(w_t^s) \Big)_{|\bar{s},\bar{t}} - \partial_s \Big(\Lambda_{S(s,t)}(v_t^s) \Big)_{|\bar{s},\bar{t}} \\
&\quad + \bar{\Lambda}_{S(\bar{s},\bar{t})}(\partial_s v_t^s - \partial_t w_t^s - [v_t^s, w_t^s]).
\end{aligned}
$$

We can now use Lemma 5.1 to conclude. QED.

Theorem 5.3 (Green's formula for smooth surfaces). *Let S be a surface in $\mathcal{M}$ satisfying the above three conditions. Let ∂S denote its boundary, defined as the union of the negatively oriented curves $S(0, \cdot)$, $S(\cdot, T)$ and the positively oriented curves $S(1, \cdot)$, $S(\cdot, 0)$. Suppose that $\mu \to ||d\Lambda_\mu||$ is also bounded on compact subsets of $\mathcal{M}$. Then*

$$\int_S d\Lambda = \int_{\partial S} \Lambda.$$

Proof: Recall that v_t^s, w_t^s and their derivatives are bounded. This, together with Equations 5.1 and 5.2, implies that the functions $(s, t) \to \Lambda_{S(s,t)}(v_t^s)$ and $(s, t) \to \Lambda_{S(s,t)}(w_t^s)$ are continuous. Hence, by Proposition 5.2, $(s, t) \to d\Lambda_{S(s,t)}(v_t^s, w_t^s)$ is Borel measurable as it is a limit of quotients of continuous functions. The fact that $\mu \to ||d\Lambda_\mu||$ is bounded on compact subsets of $\mathcal{M}$ gives that $(s, t) \to d\Lambda_{S(s,t)}(v_t^s, w_t^s)$ is bounded. The rest of the proof of this theorem is identical to that of Theorem 5.32 when we use Proposition 5.2 in place of Corollary 5.30. QED.

The regularity conditions we have imposed on S to obtain Theorem 5.3 are very strong. In particular, given an AC_2 curve σ of absolutely continuous measures in $\mathcal{M}$, it is not clear if there exists any surface S satisfying these assumptions and whose boundary is σ. It is thus difficult to use Theorem 5.3 to reach any conclusion about the de Rham cohomology of $\mathcal{M}$. Such conclusions will however be obtained in Section 5.4, based on a different version of Green's theorem.

5.2. **Regularity and differentiability of pseudo 1-forms.** The goal of this section is to introduce a regularity condition for pseudo 1-forms which guarantees differentiability. It will also ensure the validity of assumptions such as Equations 5.1 and 5.2.

Definition 5.4. Let $\bar{\Lambda}_\mu := \int_{\mathbb{R}^D} \langle \bar{A}_\mu, \cdot \rangle d\mu$ be a pseudo 1-form on $\mathcal{M}$. We say that $\bar{\Lambda}$ is *regular* if for each $\mu \in \mathcal{M}$ there exists a Borel field of $D \times D$ matrices $B_\mu \in L^\infty(\mathbb{R}^D \times \mathbb{R}^D, \mu)$ and a

function $O_\mu \in C(\mathbb{R})$ with $O_\mu(0) = 0$ such that

$$(5.7) \quad \sup_{\gamma \in \Gamma_o(\mu,\nu)} \left\{ \int_{\mathbb{R}^D \times \mathbb{R}^D} |\bar{A}_\nu(y) - \bar{A}_\mu(x) - B_\mu(x)(y-x)|^2 d\gamma(x,y) \right\}$$

$$\leq W_2^2(\mu,\nu) \min\{O_\mu(W_2(\mu,\nu)), c(\bar{\Lambda})\}^2.$$

where as usual $\Gamma_o(\mu,\nu)$ denotes the set of minimizers in Equation 2.1 and $c(\bar{\Lambda}) > 0$ is a constant independent of μ. We also assume that $||B_\mu||_\mu$ is uniformly bounded. Taking $c(\bar{\Lambda})$ large enough, there is no loss of generality in assuming that

$$(5.8) \qquad\qquad\qquad \sup_{\mu \in \mathcal{M}} ||B_\mu||_\mu \leq c(\bar{\Lambda}).$$

Let Λ be a 1-form on $\mathcal{M}$. Let $\bar{\Lambda}$ denote the associated pseudo 1-form, as in Equation 4.2. We say that Λ is *regular* if $\bar{\Lambda}$ is regular.

Remark 5.5. Some of the assumptions in Definition 5.4 could be weakened for the purposes of this paper. We make these choices simply to avoid introducing more notation and to shorten some computations.

Example 5.6. Every linear pseudo 1-form is regular. In other words, given $\bar{A} \in \mathcal{X}_c$, if we define $\bar{\Lambda}_\mu(Y) := \int_{\mathbb{R}^D} \langle \bar{A}, Y \rangle d\mu$ then $\bar{\Lambda}$ is regular. Indeed, setting $B_\mu := \nabla \bar{A}$ one can use Taylor expansion and the fact that the second derivatives of A are bounded to obtain Equation 5.7.

Remark 5.7. Even if Equation 5.7 holds for $\bar{A}_\mu$, it does not necessarily hold for $A_\mu := \pi_\mu(\bar{A})$. This implies that, in general, it is not clear what regularity properties might hold for the 1-form obtained by restricting a regular pseudo 1-form. This is true even in the simplest case where $\bar{\Lambda}$ is as in Example 5.6. The case in which $\bar{\Lambda}$ is related to Λ as in Equation 4.2 is an obvious exception: in this case, according to Definition 5.4, $\bar{\Lambda}$ is regular iff Λ is regular.

From now till the end of Section 5 we assume $\bar{\Lambda}$ is a regular pseudo 1-form on $\mathcal{M}$ and we use the notation $\bar{A}_\mu$, B_μ as in Definition 5.4.

Remark 5.8. If $\mu, \nu \in \mathcal{M}$, $X \in L^2(\mu)$, $Y \in L^2(\nu)$ and $\gamma \in \Gamma_o(\mu,\nu)$ then

$$(5.9) \quad \bar{\Lambda}_\nu(Y) - \bar{\Lambda}_\mu(X) - \int_{\mathbb{R}^D \times \mathbb{R}^D} \left(\langle \bar{A}_\mu(x), Y(y) - X(x) \rangle + \langle B_\mu(x)(y-x), Y(y) \rangle \right) d\gamma(x,y)$$

$$= \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_\nu(y) - \bar{A}_\mu(x) - B_\mu(x)(y-x), Y(y) \rangle d\gamma(x,y).$$

By Equation 5.7 and Hölder's inequality,

$$(5.10) \qquad \left| \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_\nu(y) - \bar{A}_\mu(x) - B_\mu(x)(y-x), Y(y) \rangle \right| \leq W_2(\mu,\nu)c(\bar{\Lambda}) ||Y||_\nu.$$

Similarly, Equation 5.8 and Hölder's inequality yield

$$(5.11) \qquad \left| \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle B_\mu(x)(y-x), Y(y) \rangle \right| \leq W_2(\mu,\nu)c(\bar{\Lambda}) ||Y||_\nu.$$

We use Equations 5.10 and 5.11 to obtain

$$(5.12) \quad \left| \bar{\Lambda}_\nu(Y) - \bar{\Lambda}_\mu(X) - \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_\mu(x), Y(y) - X(x) \rangle d\gamma(x,y) \right| \leq 2c(\bar{\Lambda})W_2(\mu,\nu) ||Y||_\nu.$$

Remark 5.9. Let $Y \in C_c^1(\mathbb{R}^D)$ and define $F(\mu) := \bar{\Lambda}_\mu(Y)$. Then

$$|F(\nu) - F(\mu)| \leq W_2(\nu, \mu)\Big(||\bar{A}_\nu||_\nu ||\nabla Y||_\infty + 2c(\bar{\Lambda})||Y||_\infty\Big).$$

Proof: By Hölder's inequality,

$$\left|\int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_\mu(x), Y(y) - Y(x)\rangle d\gamma(x, y)\right| \leq ||\bar{A}_\mu||_\mu ||\nabla Y||_\infty W_2(\nu, \mu).$$

We apply Remark 5.8 with $Y = X$ and we exchange the role of μ and ν to conclude the proof. QED.

Lemma 5.10. *The function*

$$\mathcal{M} \to \mathbb{R}, \quad \mu \mapsto ||\bar{A}_\mu||_\mu$$

is continuous on $\mathcal{M}$ and bounded on bounded subsets of $\mathcal{M}$. Suppose $S : [r, 1] \times [a, b] \to \mathcal{M}$ is continuous. Then

$$\sup_{(s,t) \in [r,1] \times [a,b]} ||\bar{A}_{S(s,t)}||_{S(s,t)} < \infty.$$

Proof: Fix $\mu_0 \in \mathcal{M}$. For each $\mu \in \mathcal{M}$ we choose $\gamma_\mu \in \Gamma_o(\mu_0, \mu)$. We have

$$\Big| ||\bar{A}_\mu||_\mu - ||\bar{A}_{\mu_0}||_{\mu_0} \Big| = \Big| ||\bar{A}_\mu(y)||_{\gamma_\mu} - ||\bar{A}_{\mu_0}(x)||_{\gamma_\mu} \Big| \leq ||\bar{A}_\mu(y) - \bar{A}_{\mu_0}(x)||_{\gamma_\mu}.$$

This, together with Equations 5.7 and 5.8, yields

$$\Big| ||\bar{A}_\mu||_\mu - ||\bar{A}_{\mu_0}||_{\mu_0} \Big| \leq ||B_{\mu_0}(x)(y - x)||_{\gamma_\mu} + c(\bar{\Lambda})W_2(\mu_0, \mu) \leq 2c(\bar{\Lambda})W_2(\mu_0, \mu).$$

To obtain the last inequality we have used Hölder's inequality. This proves the first claim.

Notice that $(s, t) \to ||\bar{A}_{S(s,t)}||_{S(s,t)}$ is the composition of two continuous functions and is defined on the compact set $[r, 1] \times [a, b]$. Hence it achieves its maximum. QED.

Lemma 5.11. *Let $Y \in C_c^2(\mathbb{R}^D)$ and define $F(\mu) := \bar{\Lambda}_\mu(Y)$. Then F is differentiable with gradient $\nabla_\mu F = \pi_\mu(\nabla Y^T(x)\bar{A}_\mu(x) + B_\mu^T(x)Y(x))$.*

Furthermore, assume $X \in \nabla C_c^2(\mathbb{R}^D)$ and let $\varphi_t(x) = x + tX(x) + t\bar{O}_t(x)$, where $\bar{O}_t$ is any continuous function on $\mathbb{R}^D$ such that $||\bar{O}_t||_\infty$ tends to 0 as t tends to 0. Set $\mu_t := \varphi(t, \cdot)_{\#}\mu$. Then

$$(5.13) \quad F(\mu_t) = F(\mu) + t\int_{\mathbb{R}^D} \Big[\langle \bar{A}_\mu(x), \nabla Y(x)X(x)\rangle + \langle B_\mu(x)X(x), Y(x)\rangle\Big]d\mu(x) + o(t).$$

Proof: Choose $\mu, \nu \in \mathcal{M}$ and $\gamma \in \Gamma_o(\mu, \nu)$. As in Remark 5.8,

$$\bar{\Lambda}_\nu(Y) - \bar{\Lambda}_\mu(Y) - \int_{\mathbb{R}^D \times \mathbb{R}^D} \Big(\langle \bar{A}_\mu(x), Y(y) - Y(x)\rangle + \langle B_\mu(x)(y - x), Y(y)\rangle\Big)d\gamma(x, y)$$

$$= \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_\nu(y) - \bar{A}_\mu(x) - B_\mu(x)(y - x), Y(y)\rangle d\gamma(x, y).$$

By Equation 5.7 and Hölder's inequality,

$$\left|\int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_\nu(y) - \bar{A}_\mu(x) - B_\mu(x)(y - x), Y(y)\rangle\right| \leq o(W_2(\mu, \nu)) ||Y||_\nu.$$

Since $Y \in C_c^2(\mathbb{R}^D)$ we can write $Y(y) = Y(x) + \nabla Y(x)(y - x) + R(x, y)(y - x)^2$, for some continuous field of vector-valued 2-tensors $R = R(x, y)$. In particular, R has compact support and depends on the second derivatives of Y. Then

$$\int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_\mu(x), Y(y) - Y(x) \rangle d\gamma(x, y) = \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_\mu(x), \nabla Y(x)(y - x) \rangle d\gamma(x, y)$$

$$(5.14) \qquad\qquad + \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_\mu(x), R(y - x)^2 \rangle d\gamma(x, y).$$

We now want to show that the term in Equation 5.14 is of the form $o(W_2(\mu, \nu))$ as ν tends to μ. For any $\epsilon > 0$, choose a smooth compactly supported vector field $Z = Z(x)$ such that $\|\bar{A}_\mu - Z\|_\mu < \epsilon$. Then by transposing the matrix $R(y - x)$ and using Hölder's inequality we obtain

$$\left| \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_\mu(x), R(y - x)^2 \rangle d\gamma(x, y) \right| \leq \int_{\mathbb{R}^D \times \mathbb{R}^D} |\langle (R(y - x))^T (\bar{A}_\mu(x) - Z(x)), y - x \rangle| d\gamma(x, y)$$

$$+ \int_{\mathbb{R}^D \times \mathbb{R}^D} |\langle Z(x), R(y - x)^2 \rangle| d\gamma(x, y)$$

$$\leq \epsilon \|(R(y - x))\|_\infty W_2(\mu, \nu) + \|Z\|_\infty \|R\|_\infty W_2^2(\mu, \nu).$$

Since ϵ and $\|Z\|_\infty$ are independent of ν, this gives the required estimate. Likewise,

$$\int_{\mathbb{R}^D \times \mathbb{R}^D} \langle B_\mu(x)(y - x), Y(y) \rangle d\gamma(x, y) = \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle B_\mu(x)(y - x), Y(y) - Y(x) \rangle d\gamma(x, y)$$

$$+ \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle B_\mu(x)(y - x), Y(x) \rangle d\gamma(x, y)$$

$$= \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle B_\mu(x)(y - x), Y(x) \rangle d\gamma(x, y) + o(W_2(\mu, \nu)).$$

Combining these results shows that

$$\bar{\Lambda}_\nu(Y) = \bar{\Lambda}_\mu(Y) + \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \nabla Y^T(x) \bar{A}_\mu(x) + B_\mu^T(x) Y(x), y - x \rangle d\gamma(x, y) + o(W_2(\mu, \nu)).$$

As in Definition 4.9, this proves that F is differentiable and that $\nabla_\mu F = \pi_\mu (\nabla Y^T(x) \bar{A}_\mu(x) + B_\mu^T(x) Y(x))$.

Now assume that ϕ_t is the flow of X. Notice that the curve $t \to \mu_t$ belongs to $AC_2(-r, r; \mathcal{M})$ for $r > 0$. We could choose for instance $r = 1$. Hence the curve is continuous on $[-1, 1]$. By Lemma 5.10, the composed function $t \to \|\bar{A}_{\mu_t}\|_{\mu_t}$ is also continuous. Hence its range is compact in $\mathbb{R}$, so there exists $\bar{C} > 0$ such that $\|\bar{A}_{\mu_t}\|_{\mu_t} \leq \bar{C}$ for all $t \in [-1, 1]$. We may now use Remark 4.10 to conclude.

The general case of ϕ_t as in the statement of Lemma 5.11 can be studied using analogous methods. QED.

Lemma 5.12. *Any regular pseudo 1-form $\bar{\Lambda}$ is differentiable in the sense of Definition 4.12. Furthermore, $\forall X, Y \in T_\mu \mathcal{M}$,*

$$(5.15) \qquad\qquad d\bar{\Lambda}_\mu(X, Y) = \int_{\mathbb{R}^D} \langle (B_\mu - B_\mu^T) X, Y \rangle d\mu.$$

Proof: The fact that, for each $Y \in \mathcal{X}_c$, $\bar{\Lambda}(Y)$ is a differentiable function on $\mathcal{M}$ follows from Lemma 5.11. Lemma 5.11 also gives an expression for the gradient of this function. Using this expression it is simple to check that, for $X, Y \in \nabla C_c^\infty$,

$$(5.16) \qquad X\bar{\Lambda}(Y) - Y\bar{\Lambda}(X) - \bar{\Lambda}([X,Y]) = \int_{\mathbb{R}^D} \langle (B_\mu - B_\mu^T)X, Y \rangle d\mu.$$

This proves that $\bar{\Lambda}$ is differentiable. By continuity, the same expression holds for any $X, Y \in T_\mu \mathcal{M}$.   QED.

5.3. Regular forms and absolutely continuous curves.
The goal of this section is to study the regularity and integrability properties of regular pseudo 1-forms evaluated along curves in $AC_2(a, b; \mathcal{M})$.

Lemma 5.13. *Assume $\{\mu_\epsilon\}_{\epsilon \in E} \subset \mathcal{M}$ and $v_\epsilon \in L^2(\mu_\epsilon)$ are such that $C := \sup_{\epsilon \in E} ||v_\epsilon||_{L^2(\mu_\epsilon)}$ is finite. Assume $\{\mu_\epsilon\}_{\epsilon \in E}$ converges to μ in $\mathcal{M}$ as ϵ tends to 0 and that there exists $v \in L^2(\mu)$ such that $\{v_\epsilon \mu_\epsilon\}_{\epsilon \in E}$ converges weak-* to $v\mu$ as $\epsilon \to 0$. If $\gamma_\epsilon \in \Gamma_o(\mu, \mu_\epsilon)$ then $\lim_{\epsilon \to 0} a_\epsilon = 0$, where $a_\epsilon = \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_\mu(x), v_\epsilon(y) - v(x) \rangle d\gamma_\epsilon(x, y)$.*

Proof: It is easy to obtain that $||v||_{L^2(\mu)} \leq C$. Let $\gamma_\epsilon \in \Gamma_o(\mu, \mu_\epsilon)$ and $\xi \in \mathcal{X}_c$. Then there exists a bounded function $C_\xi \in C(\mathbb{R}^D \times \mathbb{R}^D)$ and a real number M such that

$$(5.17) \qquad \xi(x) - \xi(y) = \nabla\xi(y)(x - y) + |x - y|^2 C_\xi(x, y), \quad |C_\xi(x, y)| \leq M,$$

for $x, y \in \mathbb{R}^D$. We use the first equality in Equation 5.17 to obtain that

$$\langle \bar{A}_\mu(x), v_\epsilon(y) - v(x) \rangle = \langle \bar{A}_\mu(x) - \xi(x), v_\epsilon(y) - v(x) \rangle + \langle \xi(y), v_\epsilon(y) \rangle - \langle \xi(x), v(x) \rangle$$
$$+ \langle \nabla\xi(y)(x - y) + |x - y|^2 C_\xi(x, y), v_\epsilon(y) \rangle.$$

Hence,

$$|a_\epsilon| \leq ||\bar{A}_\mu(x) - \xi(x)||_{L^2(\gamma_\epsilon)} ||v_\epsilon(y) - v(x)||_{L^2(\gamma_\epsilon)} + b_\epsilon$$
$$(5.18) \qquad + \left| \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle (\nabla\xi(y)(x - y) + |x - y|^2 C_\xi(x, y)), v_\epsilon(y) \rangle d\gamma_\epsilon(x, y) \right|.$$

Above, we have set $b_\epsilon := \left| \int_{\mathbb{R}^D \times \mathbb{R}^D} (\langle \xi(y), v_\epsilon(y) \rangle - \langle \xi(x), v(x) \rangle) d\gamma_\epsilon(x, y) \right|$. By the second inequality in Equation 5.17 and by Equation 5.18,

$$(5.19) \qquad |a_\epsilon| \leq 2C||\bar{A}_\mu - \xi||_{L^2(\mu)} + b_\epsilon + ||\nabla\xi||_\infty W_2(\mu, \mu_\epsilon) + MW_2^2(\mu, \mu_\epsilon).$$

By assumption $\{W_2(\mu, \mu_\epsilon)\}_{\epsilon \in E}$ tends to 0 and $\{b_\epsilon\}_{\epsilon \in E}$ tends to 0 as ϵ tends to 0. These facts, together with Equation 5.19, yield $\limsup_{\epsilon \to 0} |a_\epsilon| \leq 2C||\bar{A}_\mu - \xi||_{L^2(\mu)}$ for arbitrary $\xi \in \mathcal{X}_c$. We use that $\mathcal{X}_c$ is dense in $L^2(\mu)$ to conclude that $\lim_{\epsilon \to 0} a_\epsilon = 0$.   QED.

Corollary 5.14. *Assume $\{\mu_\epsilon\}_{\epsilon \in E} \subset \mathcal{M}$, μ, $v_\epsilon \in L^2(\mu_\epsilon)$ and v satisfy the assumptions of Lemma 5.13. Then $\lim_{\epsilon \to 0} \bar{\Lambda}_{\mu_\epsilon}(v_\epsilon) = \bar{\Lambda}_\mu(v)$.*

Proof: Let $\gamma_\epsilon \in \Gamma_o(\mu, \mu_\epsilon)$. Observe that

$$\langle \bar{A}_{\mu_\epsilon}(y), v_\epsilon(y) \rangle - \langle \bar{A}_\mu(x), v(x) \rangle = \langle \bar{A}_\mu(x), v_\epsilon(y) - v(x) \rangle + \langle B_\mu(x)(y - x), v_\epsilon(y) \rangle$$
$$(5.20) \qquad + \left\langle \bar{A}_{\mu_\epsilon}(y) - \bar{A}_\mu(x) - B_\mu(x)(y - x), v_\epsilon(y) \right\rangle.$$

We now integrate Equation 5.20 over $\mathbb{R}^D \times \mathbb{R}^D$ and use Equations 5.7, 5.8 and the fact that $\gamma_\epsilon \in \Gamma_o(\mu, \mu_\epsilon)$. We obtain

$$|\bar{\Lambda}_{\mu_\epsilon}(v_\epsilon) - \bar{\Lambda}_\mu(v)| \le |a_\epsilon| + ||B_\mu||_{L^\infty(\mu)} W_2(\mu, \mu_\epsilon) ||v_\epsilon||_{\mu_\epsilon} + o(W_2(\mu, \mu_\epsilon)) ||v_\epsilon||_{\mu_\epsilon}$$
$$\text{(5.21)} \qquad\qquad \le |a_\epsilon| + C ||B_\mu||_{L^\infty(\mu)} W_2(\mu, \mu_\epsilon) + C\, o(W_2(\mu, \mu_\epsilon)).$$

Letting ϵ tend to 0 in Equation 5.21 we conclude the proof of the corollary.   QED.

Lemma 5.15 (continuity of $\bar{\Lambda}_{\sigma_t}(X_t)$). *Suppose $\sigma \in AC_2(a, b; \mathcal{M})$. If $X \in C((a, b) \times \mathbb{R}^D, \mathbb{R}^D)$ then $\lambda(t) := \bar{\Lambda}_{\sigma_t}(X_t)$ is continuous on (a, b).*

Proof: Fix $t \in (a, b)$ so that t belongs to the interior of a compact set $K^* \subset (a, b)$. Let $\varphi \in C_c(\mathbb{R}^D, \mathbb{R}^D)$ and denote by K a compact set containing its support. Observe that X is uniformly continuous on $K^* \times K$ so
(5.22)

$$\limsup_{h \to 0} |\int_{\mathbb{R}^D} \langle \varphi(x), X_{t+h}(x) - X_t(x)\rangle d\sigma_{t+h}(x)| \le \limsup_{h \to 0} ||\varphi||_\infty \sup_{x \in K} |X_{t+h}(x) - X_t(x)| = 0.$$

Since $\langle X_t, \varphi \rangle \in C_c$ and σ is continuous at t by Remark 2.11, we also see that

$$\text{(5.23)} \qquad \lim_{h \to 0} \int_{\mathbb{R}^D} \langle \varphi(x), X_t(x)\rangle d\sigma_{t+h}(x) = \int_{\mathbb{R}^D} \langle \varphi(x), X_t(x)\rangle d\sigma_t(x).$$

Since $\varphi \in C_c(\mathbb{R}^D, \mathbb{R}^D)$ is arbitrary, Equations 5.22 and 5.23 show that $\{X_{t+h}\sigma_{t+h}\}_{h>0}$ converges weak-$*$ to $\sigma_t X_t$ as h tends to zero. Corollary 5.14 yields that λ is continuous at t.

QED.

Remark 5.16. Using the techniques of Lemma 5.29 one could further prove that if $\sigma \in AC_2(a, b; \mathcal{M})$ and X is sufficiently regular then $\lambda(t) := \bar{\Lambda}_{\sigma_t}(X_t)$ is Lipschitz and $\mathcal{L}^1$-almost everywhere differentiable.

We now assume that $\eta_D^\epsilon \in C^\infty(\mathbb{R}^D)$ is a mollifier : $\eta_D^\epsilon(x) = 1/\epsilon^D \eta(x/\epsilon)$, for some bounded symmetric function $\eta \in C^\infty(\mathbb{R}^D)$ whose derivatives of all orders are bounded. We also impose that $\eta > 0$, $\int_{\mathbb{R}^D} |x|^2 \eta(x) dx < \infty$ and $\int_{\mathbb{R}^D} \eta = 1$. We fix $\mu \in \mathcal{M}$ and define $f^\epsilon(x) := \int_{\mathbb{R}^D} \eta_D^\epsilon(x - y) d\mu(y)$. Observe that $f^\epsilon \in C^\infty(\mathbb{R}^D)$ is bounded, all its derivatives are bounded and $\int_{\mathbb{R}^D} f^\epsilon = 1$.

We suppose that $\eta_1^\epsilon \in C^\infty(\mathbb{R})$ is a standard mollifier: $\eta_1^\epsilon(t) = 1/\epsilon \eta_1(t/\epsilon)$, for some bounded symmetric function $\eta_1 \in C^\infty(\mathbb{R})$ which is positive on $(-1, 1)$ and vanishes outside $(-1, 1)$. We also impose that $\int_{\mathbb{R}} \eta_1 = 1$ and assume that $|\epsilon| < 1$.

Suppose $\sigma \in AC_2(a, b; \mathcal{M})$ and $v : (a, b) \times \mathbb{R}^D \to \mathbb{R}^D$ is a velocity associated to σ so that $t \to ||v_t||_{\sigma_t} \in L^\infty(a, b)$. Suppose that for each $t \in (a, b)$ there exists $\rho_t > 0$ such that $\sigma_t = \rho_t \mathcal{L}^D$.

We can extend σ and v in time on an interval larger than $[a, b]$. For instance, set $\tilde{\sigma}_t = \sigma_a$ for $t \in (a - 1, a)$ and set $\tilde{\sigma}_t = \sigma_b$ for $t \in (b, b + 1)$. Observe that $\tilde{\sigma} \in AC_2(a - 1, b + 1; \mathcal{M})$ and we have a velocity $\tilde{v}$ associated to $\tilde{\sigma}$ such that $\tilde{v}_t = v_t$ for $t \in [a, b]$. We can choose $\tilde{v}$ such that $||\tilde{v}_t||_{\tilde{\sigma}_t}^2 = 0$ for t outside (a, b). In particular, $\int_{a-1}^{b-1} ||\tilde{v}_t||_{\tilde{\sigma}_t}^2 dt = \int_a^b ||v_t||_{\sigma_t}^2 dt$. In the sequel we won't distinguish between $\sigma, \tilde{\sigma}$ on the one hand and $v, \tilde{v}$ on the other hand. This

extension becomes useful when we try to define ρ_t^ϵ as it appears in Equation 5.24. The new density functions are meaningful if we substitute σ by $\tilde{\sigma}$ and impose that $\epsilon \in (0,1)$.

For $\epsilon \in (0,1)$, set

$$(5.24) \quad \rho_t^\epsilon(x) := \int_{\mathbb{R}} \eta_1^\epsilon(t-\tau)\rho_\tau(x)d\tau, \;\; \sigma_t^\epsilon := \rho_t^\epsilon \mathcal{L}^D, \;\; \rho_t^\epsilon(x)v_t^\epsilon(x) := \int_{\mathbb{R}} \eta_1^\epsilon(t-\tau)\rho_\tau(x)v_\tau(x)d\tau.$$

Note that $\rho_t^\epsilon(x) > 0$ for all $t \in (a,b)$ and $x \in \mathbb{R}^D$ and ρ_t^ϵ is a probability density. Also, $v^\epsilon : (a,b) \times \mathbb{R}^D \to \mathbb{R}^D$ is a velocity associated to σ^ϵ. In the sequel we set

$$C^2 := \int_{\mathbb{R}^D} |x|^2 \eta(x)dx, \quad C_1 = \int_{\mathbb{R}} \eta_1(\tau)\tau d\tau, \quad C_v := \sup_{\tau \in (a-1, b+1)} ||v_\tau||_{\sigma_\tau}.$$

Lemma 5.17. *We assume that for each $t \in (a,b)$ there exists $\rho_t > 0$ such that $\sigma_t = \rho_t \mathcal{L}^D$. Then $\sigma^\epsilon \in AC_2(a,b;\mathcal{M})$. For $a < s < t < b$,*

$$(i) \;\; W_2(\mu, f^\epsilon \mathcal{L}^D) \le \epsilon C, \quad (ii) \;\; ||v_t^\epsilon||_{\sigma_t^\epsilon} \le C_v \quad and \quad (iii) \;\; W_2(\sigma_t^\epsilon, \sigma_t) \le \epsilon C_1 C_v.$$

Proof: We denote by $\mathcal{U}$ the set of pairs (u,v) such that $u,v \in C(\mathbb{R}^D)$ are bounded and $u(x) + v(y) \le |x-y|^2$ for all $x, y \in \mathbb{R}^D$. Fix $(u,v) \in \mathcal{U}$. By Fubini's theorem one gets the well known identity

$$(5.25) \qquad \int_{\mathbb{R}^D} u(x)f^\epsilon(x)dx = \int_{\mathbb{R}^D} d\mu(y) \int_{\mathbb{R}^D} u(x)\eta_\epsilon(x-y)dx.$$

Since $v(y) = \int_{\mathbb{R}^D} v(y)\eta_\epsilon(x-y)dx$, Equation 5.25 yields that

$$\int_{\mathbb{R}^D} u(x)f^\epsilon(x)dx + \int_{\mathbb{R}^D} v(y)d\mu(y) = \int_{\mathbb{R}^D} d\mu(y) \int_{\mathbb{R}^D} \eta_\epsilon(x-y)\big(u(x) + v(y)\big)dx$$

$$(5.26) \qquad\qquad\qquad \le \int_{\mathbb{R}^D} d\mu(y) \int_{\mathbb{R}^D} \eta_\epsilon(x-y)|x-y|^2 dx$$

$$\qquad\qquad\qquad = \int_{\mathbb{R}^D} d\mu(y) \int_{\mathbb{R}^D} \frac{1}{\epsilon^D}\eta(\frac{z}{\epsilon})|z|^2 dz = C^2 \epsilon^2.$$

To obtain Equation 5.26 we have used that $(u,v) \in \mathcal{U}$. We have proven that $\int_{\mathbb{R}^D} u(x)f^\epsilon(x)\,dx + \int_{\mathbb{R}^D} v(y)\,d\mu(y) \le C^2\epsilon^2$ for arbitrary $(u,v) \in \mathcal{U}$. Thanks to the dual formulation of the Wasserstein distance Equation 2.2, we conclude the proof of (i).

Notice that for each $t \in (a,b)$ and $x \in \mathbb{R}^D$, $\eta_1^\epsilon(t-\tau)\rho_\tau(x)/\rho_t^\epsilon(x)$ is a probability density on $\mathbb{R}$. Hence, by Jensen's inequality,

$$|v_t^\epsilon(x)|^2 = \left|1/\rho_t^\epsilon(x) \int_{\mathbb{R}} \eta_1^\epsilon(t-\tau)\rho_\tau(x)v_\tau(x)d\tau\right|^2 \le 1/\rho_t^\epsilon(x) \int_{\mathbb{R}} \eta_1^\epsilon(t-\tau)\rho_\tau(x)|v_\tau(x)|^2 d\tau.$$

We multiply both sides of the previous inequality by $\rho_t^\epsilon(x)$. We then integrate the subsequent inequality over $\mathbb{R}^D$ and use Fubini's theorem to conclude the proof of (ii).

We use (ii) and Remark 2.11 (i) to obtain that $\sigma^\epsilon \in AC_2(a,b;\mathcal{M})$. We have

$$\int_{\mathbb{R}^D} u(x)d\sigma_t^\epsilon(x) = \int_{\mathbb{R}^D} u(x)dx \int_{\mathbb{R}} \eta_1^\epsilon(\tau)\rho_{t-\tau}(x)d\tau = \int_{\mathbb{R}} \eta_1^\epsilon(\tau)d\tau \int_{\mathbb{R}^D} u(x)d\sigma_{t-\tau}(x).$$

Hence, using that $v(y) = \int_{\mathbb{R}} \eta_1^\epsilon(\tau) v(y) d\tau$, we obtain

$$\int_{\mathbb{R}^D} u(x) d\sigma_t^\epsilon(x) + \int_{\mathbb{R}^D} v(y) d\sigma_t(y) = \int_{\mathbb{R}} \eta_1^\epsilon(\tau) d\tau \left(\int_{\mathbb{R}^D} u d\sigma_{t-\tau} + \int_{\mathbb{R}^D} v d\sigma_t \right)$$

$$(5.27) \qquad\qquad\qquad\qquad \leq \int_{\mathbb{R}} \eta_1^\epsilon(\tau) W_2^2(\sigma_{t-\tau}, \sigma_t) d\tau$$

$$(5.28) \qquad\qquad\qquad\qquad \leq \int_{\mathbb{R}} \eta_1^\epsilon(\tau) \tau^2 C_v^2 d\tau = \epsilon^2 C_1 C_v^2.$$

To obtain Equation 5.27 we have used the dual formulation of the Wasserstein distance Equation 2.2 and the fact that $(u, v) \in \mathcal{U}$. We have used Remark 2.11 to obtain Equation 5.28. Since $\int_{\mathbb{R}^D} u \, d\sigma_t^\epsilon + \int_{\mathbb{R}^D} v \, d\sigma_t \leq \epsilon C C_v$ for arbitrary $(u, v) \in \mathcal{U}$, we conclude that (iii) holds.

$$\text{QED.}$$

Remark 5.18. Assume that for each $t \in (a, b)$ there exists $\rho_t > 0$ such that $\sigma_t = \rho_t \mathcal{L}^D$. Let $\phi \in C_c(\mathbb{R}^D)$. Setting $I_\phi(t) := \int_{\mathbb{R}^D} \langle \phi, v_t \rangle \rho_t d\mathcal{L}^D$, we have

$$(5.29) \qquad\qquad | \int_{\mathbb{R}^D} \langle \phi, v_t^\epsilon \rangle \rho_t^\epsilon d\mathcal{L}^D | = |\eta_1^\epsilon * I_\phi(t)| \leq ||\phi||_\infty C_v.$$

Corollary 5.19. *Suppose that for each $t \in (a, b)$ there exists $\rho_t > 0$ such that $\sigma_t = \rho_t \mathcal{L}^D$. Then, for each $t \in [a, b]$, $\{\sigma_t^\epsilon\}_{\epsilon > 0}$ converges to σ_t in $\mathcal{M}$ as ϵ tends to zero. For $\mathcal{L}^1$-almost every $t \in [a, b]$, $\{\sigma_t^\epsilon v_t^\epsilon\}_{\epsilon > 0}$ converges weak-$*$ to $\sigma_t v_t$ as ϵ tends to zero.*

Proof: By Lemma 5.17 (iii), $\{\sigma_t^\epsilon\}_{\epsilon > 0}$ converges to σ_t in $\mathcal{M}$ as ϵ tends to zero.

Let $\mathcal{C}$ be a countable family in $C_c(\mathbb{R}^D)$. For each $\phi \in C_c(\mathbb{R}^D)$, the set of Lebesgue points of I_ϕ is a set of full measure in $[a, b]$. For these points $\eta_1^\epsilon * I_\phi(t)$ tends to $I_\phi(t)$ as ϵ tends to zero. Thus there is a set S of full measure in $[a, b]$ such that for all $\phi \in \mathcal{C}$ and all $t \in S$, $\eta_1^\epsilon * I_\phi(t)$ tends to $I_\phi(t)$ as ϵ tends to zero. Fix $\varphi \in C_c(\mathbb{R}^D)$ and choose $\delta > 0$ arbitrary. Let $\phi \in \mathcal{C}$ be such that $||\varphi - \phi||_\infty \leq \delta$. Note that

$$|\eta_1^\epsilon * I_\varphi(t) - I_\varphi(t)| \leq |\eta_1^\epsilon * I_\phi(t) - I_\phi(t)| + |\eta_1^\epsilon * I_{\phi - \varphi}(t)| + |I_{\phi - \varphi}(t)|.$$

We use the inequality in Equation 5.29 to conclude that

$$|\eta_1^\epsilon * I_\varphi(t) - I_\varphi(t)| \leq |\eta_1^\epsilon * I_\phi(t) - I_\phi(t)| + 2\delta C_v.$$

If $t \in S$, the previous inequality gives that $\limsup_{\epsilon \to 0} |\eta_1^\epsilon * I_\varphi(t) - I_\varphi(t)| \leq 2\delta C_v$. Since $\delta > 0$ is arbitrary we conclude that $\lim_{\epsilon \to 0} |\eta_1^\epsilon * I_\varphi(t) - I_\varphi(t)| = 0$. $\qquad$ QED.

Corollary 5.20. *Suppose that $\sigma \in AC_2(a, b; \mathcal{M})$ for all $a < b$, v is a velocity associated to σ and $C := \sup_{t \in [a, b]} ||v_t||_{\sigma_t} < \infty$. Define*

$$f_t^r(x) := \int_{\mathbb{R}^D} \eta_D^r(x - y) d\sigma_t(y), \quad \sigma_t^r := f_t^r \mathcal{L}^D, \quad f_t^r(x) v_t^r(x) := \int_{\mathbb{R}^D} \eta_D^r(x - y) v_t(y) d\sigma_t(y).$$

As in Equation 5.24, for $0 < \epsilon < 1$ we define

$$\rho_t^{\epsilon, r}(x) := \int_{\mathbb{R}} \eta_1^\epsilon(t - \tau) f_\tau^r(x) d\tau, \quad \sigma_t^{\epsilon, r} := \rho^{\epsilon, r} \mathcal{L}^D, \quad \rho_t^{\epsilon, r}(x) v_t^{\epsilon, r}(x) := \int_{\mathbb{R}} \eta_1^\epsilon(t - \tau) f_\tau^r(x) v_\tau^r(x) d\tau.$$

Then:

(i) v^r is a velocity associated to σ^r and, for each $t \in (a, b)$, $\{\sigma_t^r\}_r$ converges to σ_t in $\mathcal{M}$ as r tends to zero. For $\mathcal{L}^1$-almost every $t \in (a, b)$, $||v_t^r||_{\sigma_t^r} \leq C$ and $\{v_t^r \sigma_t^r\}_{r>0}$ converges weak-$*$ to $v_t \sigma_t$ as r tends to zero.

(ii) $v^{\epsilon,r}$ is a velocity associated to $\sigma^{\epsilon,r}$ and, for each $t \in (a, b)$, $\{\sigma_t^{\epsilon,r}\}_\epsilon$ converges to σ_t^r in $\mathcal{M}$ as ϵ tends to zero. For every $t \in (a, b)$, $||v_t^{\epsilon,r}||_{\sigma_t^{\epsilon,r}} \leq C$ while for $\mathcal{L}^1$-almost every $t \in (a, b)$, $\{v_t^{\epsilon,r} \sigma_t^{\epsilon,r}\}_{r>0}$ converges weak-$*$ to $v_t^r \sigma_t^r$ as ϵ tends to zero.

(iii) The function $t \to \bar{\Lambda}_{\sigma_t^{\epsilon,r}}(v_t^{\epsilon,r})$ is continuous while $t \to \bar{\Lambda}_{\sigma_t}(v_t)$ is measurable on (a, b).

(iv) Suppose in addition that σ is time-periodic in the sense that $\sigma_a = \sigma_b$. Then $\sigma_a^r = \sigma_b^r$.

Proof: It is well known that $||v_t^r||_{\sigma_t^r} \leq ||v_t||_{\sigma_t} \leq C$ (cf. [5] Lemma 8.1.10) so, by Remark 2.11 (i), $\sigma \in AC_2(a, b; \mathcal{M})$. One can readily check that v^r is a velocity associated to σ^r. Lemma 5.17 shows that, for each $t \in (a, b)$, $\{\sigma_t^r\}_r$ converges to σ_t in $\mathcal{M}$ as r tends to zero. Let $\varphi \in C_c(\mathbb{R}^D, \mathbb{R}^D)$. Set $\varphi^r := \eta_D^r * \varphi$. Since $\{\varphi^r\}_{r>0}$ converges uniformly to φ,

$$\lim_{r \to 0} \int_{\mathbb{R}^D} \langle \varphi, v_t^r \rangle d\sigma_t^r = \int_{\mathbb{R}^D} \langle v_t, \varphi \rangle d\sigma_t.$$

Thus $\{v_t^r \sigma_t^r\}_{r>0}$ converges weak-$*$ to $v_t \sigma_t$ as r tends to zero. This proves (i).

We next fix $r > 0$. For a moment we won't display the dependence in r. For instance we write v^ϵ instead of $v_t^{\epsilon,r}$ as in Equation 5.24. Notice that $\rho^\epsilon \in C^1([a, b] \times \mathbb{R}^D)$, $\rho^\epsilon > 0$ and ρ_t^ϵ is a probability density. Also $v_t^\epsilon \in C^1([a, b] \times \mathbb{R}^D, \mathbb{R}^D)$ and v^ϵ is a velocity associated to σ^ϵ. Fix $t \in [\bar{a}, \bar{b}] \subset (a, b)$. Lemma 5.17 gives that $||v_t^\epsilon||_{\sigma_t^\epsilon} \leq C$ for all $\epsilon > 0$ small enough. By Corollary 5.19, $\{v_t^\epsilon \sigma_t^\epsilon\}_{\epsilon>0}$ converges weak-$*$ to $v_t \sigma_t$ as ϵ tends to zero. This proves (ii).

By Lemma 5.15, $t \to \bar{\Lambda}_{\sigma_t^\epsilon}(v_t^\epsilon)$ is continuous in (a, b). Hence by (ii) $t \to \bar{\Lambda}_{\sigma_t^r}(v_t^r)$ is measurable as a pointwise limit of measurable functions. We then use (i) to conclude that $t \to \bar{\Lambda}_{\sigma_t}(v_t)$ is measurable as a pointwise limit of measurable functions. This proves (iii). The proof of (iv) is straightforward.

QED.

We can now prove that regular pseudo 1-forms can be integrated along absolutely continuous curves, as follows.

Corollary 5.21. Let $\sigma \in AC_2(a, b; \mathcal{M})$ and let v be a velocity associated to σ. Suppose $t \to ||v_t||_{\sigma_t}$ is square integrable on (a, b). Then $t \to \bar{\Lambda}_{\sigma_t}(v_t)$ is measurable and square integrable on (a, b).

Proof: Let $\bar{\sigma}$ be the reparametrization of σ as introduced in Remark 2.13 and let $\bar{v}$ be the associated velocity. By Corollary 5.20 (iii), because $\sup_{s \in [0,L]} ||\bar{v}_s||_{\bar{\sigma}_s} \leq 1$, we have that $s \to \bar{\Lambda}_{\bar{\sigma}_s}(\bar{v}_s)$ is measurable. But $\bar{\Lambda}_{\sigma_t}(v_t) = \dot{S}(t)\bar{\Lambda}_{\bar{\sigma}_{S(t)}}(\bar{v}_{S(t)})$. Thus $t \to \bar{\Lambda}_{\sigma_t}(v_t)$ is measurable.

By Corollary 5.10 there exists a constant C_σ independent of t such that $||\bar{A}_{\sigma_t}||_{\sigma_t} \leq C_\sigma$ for all $t \in [a, b]$. Thus

$$|\bar{\Lambda}_{\sigma_t}(v_t)| = \left| \int_{\mathbb{R}^D} \langle \bar{A}_{\sigma_t}, v_t \rangle d\sigma_t \right| \leq ||\bar{A}_{\sigma_t}||_{\sigma_t} ||v_t||_{\sigma_t} \leq C_\sigma ||v_t||_{\sigma_t}.$$

Since $t \to ||v_t||_{\sigma_t}$ is square integrable, the previous inequality yields the proof. QED.

Corollary 5.22. Suppose $\{\sigma^r\}_{0 \leq r \leq c} \subset AC_2(a, b; \mathcal{M})$, v^r is a velocity associated to σ^r and $\infty > C := \sup_{(t,r) \in E} ||v_t^r||_{\sigma_t^r}$ where $E := [a, b] \times [0, c]$. Suppose that, for $\mathcal{L}^1$-almost every $t \in (a, b)$, $\{v_t^r \sigma_t^r\}_{r>0}$ converges weak-$*$ to $v_t \sigma_t$ and $\{\sigma_t^r\}_{r>0}$ converges in $\mathcal{M}$ to σ_t as r tends

to zero. If $(t, r) \to \sigma_t^r$ is continuous at every $(t, 0) \in [a, b] \times \{0\}$ then $\lim_{r \to 0} \int_a^b \bar{\Lambda}_{\sigma^r}(v^r) dt = \int_a^b \bar{\Lambda}_\sigma(v) dt$. Here we have set $\sigma_t := \sigma_t^0$.

Proof: By Lemma 5.10 we may assume without loss of generality that $||\bar{A}_{\sigma_t^r}||_{\sigma_t^r}$ is bounded on E by a constant $\bar{C}$ independent of $(t, r) \in E$. We obtain

(5.30) $$\sup_{(t,r) \in E} |\bar{\Lambda}_{\sigma_t^r}(v_t^r)| \leq \sup_{(t,r) \in E} ||\bar{A}_{\sigma_t^r}||_{\sigma_t^r} ||v_t^r||_{\sigma_t^r} \leq \bar{C} C.$$

Corollary 5.14 ensures that $\lim_{r \to 0} \bar{\Lambda}_{\sigma_t^r}(v_t^r) = \bar{\Lambda}_{\sigma_t}(v_t)$ for $\mathcal{L}^1$-almost every $t \in [a, b]$. This, together with Equation 5.30 shows that, as r tends to 0, the sequence of functions $t \to \bar{\Lambda}_{\sigma_t^r}(v_t^r)$ converges to the function $t \to \bar{\Lambda}_{\sigma_t}(v_t)$ in $L^1(a, b)$. This proves the corollary.   QED.

Definition 5.23. Let $\sigma \in AC_2(a, b; \mathcal{M})$ and let v be a velocity associated to σ. Suppose $t \to ||v_t||_{\sigma_t}$ is square integrable on (a, b). By Corollary 5.21, $t \to \bar{\Lambda}_{\sigma_t}(v_t)$ is also square integrable on (a, b). It is thus meaningful to calculate the integral $\int_a^b \bar{\Lambda}_{\sigma_t}(v_t) dt$.

We will call $\int_a^b \bar{\Lambda}_{\sigma_t}(v_t) dt$ the *integral of $\bar{\Lambda}$ along (σ, v)*. When v is the velocity of minimal norm we will write this simply as $\int_\sigma \bar{\Lambda}$ and call it the *integral of $\bar{\Lambda}$ along σ*.

Remark 5.24. Suppose that $r : [c, d] \to [a, b]$ is invertible and Lipschitz. Define $\bar{\sigma}_s = \sigma_{r(s)}$. Then $\bar{\sigma} \in AC_2(c, d; \mathcal{M})$ and $\bar{v}_s(x) = \dot{r}(s) v_{r(s)}(x)$ is a velocity for $\bar{\sigma}$. Furthermore, $\int_c^d \bar{\Lambda}_{\bar{\sigma}_t}(\bar{v}_t) dt = \int_a^b \bar{\Lambda}_{\sigma_t}(v_t) dt$.

Proof: Let $\beta \in L^2(a, b)$ be as in Definition 2.10. Then

$$W_2(\sigma_{r(s+h)}, \sigma_{r(s)}) \leq \left| \int_{r(s)}^{r(s+h)} \beta(t) dt \right| = \left| \int_s^{s+h} \bar{\beta}(\tau) d\tau \right| \quad \text{where} \quad \bar{\beta}(s) := |\dot{r}(s)| \beta(r(s)).$$

Because $\bar{\beta} \in L^2(c, d)$ we conclude that $\bar{\sigma} \in AC_2(c, d; \mathcal{M})$. Direct computations give that, for $\mathcal{L}^1 - $ a.e. $s \in (c, d)$,

$$\lim_{h \to 0} W_2(\sigma_{r(s+h)}, \sigma_{r(s)})/|h| = |\dot{r}(s)| \, |\sigma'|(r(s)).$$

Thus $|\bar{\sigma}'|(s) = |\dot{r}(s)| \, |\sigma'|(r(s))$. Let $\phi \in C_c^\infty(\mathbb{R}^D)$ and let v be a velocity for σ (see Proposition 2.12). The chain rule shows that, in the sense of distributions,

$$\frac{d}{ds} \int_{\mathbb{R}^D} \phi \, d\sigma_{r(s)} = \dot{r}(s) \langle \nabla \phi, v_{r(s)} \rangle_{\sigma_{r(s)}} = \langle \nabla \phi, \bar{v}_s \rangle_{\bar{\sigma}_s},$$

where $\bar{v}_s(x) = \dot{r}(s) v_{r(s)}(x)$. Thus $\bar{v}$ is a velocity for $\bar{\sigma}$. Using the linearity of $\bar{\Lambda}$ we have

$$\int_c^d \bar{\Lambda}_{\bar{\sigma}_s}(\bar{v}_s) ds = \int_c^d \dot{r}(s) \bar{\Lambda}_{\sigma_{r(s)}}(v_{r(s)}) ds = \int_a^b \bar{\Lambda}_{\sigma_t}(v_t) dt.$$

This concludes the proof.   QED.

5.4. **Green's formula for annuli.** Let $\sigma \in AC_2(a, b; \mathcal{M})$. Given $r \in (0, 1)$ and $s \in [r, 1]$, let $D_s : \mathbb{R}^D \to \mathbb{R}^D$ denote the map defined by $D_s(x) := sx$. Using this map we can canonically associate to σ the surface

$$(5.31) \qquad \sigma(s, t) : [r, 1] \times (a, b) \to \mathcal{M}, \quad \sigma(s, t) = \sigma_t^s := D_{s\#}\sigma_t.$$

We call such a surface an *annulus*. We now want to study its properties.

Let v denote the velocity of σ of minimal norm. Set

$$w(s, t, \cdot) = w_t^s(x) := \frac{x}{s} = D_s^{-1}(x), \quad v(s, t, \cdot) = v_t^s := D_{s*}v_t.$$

According to Lemma 4.2, for each $s \in [r, 1]$, $\sigma(s, \cdot) \in AC_2(a, b; \mathcal{M})$ admits $v(s, \cdot)$ as a velocity. For each t and $\phi \in C_c^\infty(\mathbb{R}^D)$, in the sense of distributions,

$$\frac{d}{ds} \int_{\mathbb{R}^D} \phi \, d\sigma_t^s = \frac{d}{ds} \int_{\mathbb{R}^D} \phi(sx) \, d\sigma_t(x) = \int_{\mathbb{R}^D} d\phi(sx)(x) \, d\sigma_t(x) = \int_{\mathbb{R}^D} d\phi(w_t^s) \, d\sigma_t^s.$$

Thus $w(\cdot, t)$ is a velocity for $\sigma(\cdot, t)$.

We assume that

$$||\sigma'||_\infty := \sup_{t \in [a,b]} ||v_t||_{\sigma_t} < \infty.$$

By Remark 2.11,

$$c_\sigma^0 := \sup_{t \in [a,b]} W_2(\sigma_t, \delta_0) < \infty.$$

By the fact that $D_{s\#}\sigma_t = \sigma_t^s$ we have

$$(5.32) \qquad W_2^2(\sigma_t^s, \delta_0) = s^2 W_2^2(\sigma_t, \delta_0) \le s^2 c_\sigma^0 \le \bar{C}_\sigma,$$

where we are free to choose $\bar{C}_\sigma$ to be any constant greater than c_σ^0.

Remark 5.25. Note that $(1 + h/s)Id$ pushes σ_t^s forward to σ_t^{s+h} and is the gradient of a convex function. Thus

$$\gamma^h := \Big(Id \times (1 + h/s)Id \Big)_\# \sigma_t^s \in \Gamma_o(\sigma_t^s, \sigma_t^{s+h}).$$

For γ^h-almost every $(x, y) \in \mathbb{R}^D \times \mathbb{R}^D$ we have $y = (1 + h/s)x$, so

$$(5.33) \qquad v_t^{s+h}(y) = (s + h)v_t\left(\frac{y}{s+h}\right) = (1 + \frac{h}{s})v_t^s\left(\frac{sy}{s+h}\right) = (1 + \frac{h}{s})v_t^s(x).$$

Using the definition of σ_t^s and v_t^s we obtain the identities

$$(5.34) \qquad ||Id||_{\sigma_t^s} = s||Id||_{\sigma_t} \le s\bar{C}_\sigma, \quad ||v_t^s||_{\sigma_t^s} = s||v_t||_{\sigma_t} \le s||\sigma'||_\infty.$$

We use the first identity in Equation 5.34 and the fact that $(1 + h/s)Id$ pushes σ_t^s forward to σ_t^{s+h} to obtain

$$(5.35) \qquad W_2^2(\sigma_t^s, \sigma_t^{s+h}) = \frac{h^2}{s^2}||Id||_{\sigma_t^s}^2 = h^2||Id||_{\sigma_t}^2 = h^2 W_2^2(\sigma_t, \delta_0) \le h^2 \bar{C}_\sigma^2.$$

Lemma 5.26. *There exists a constant $C_\sigma(r)$ depending only on σ and r such that $||\bar{A}_{\sigma_t^s}||_{\sigma_t^s} \le C_\sigma(r)$, for all $(s, t) \in [r, 1] \times [a, b]$.*

Proof: By Remark 2.11 (i), $\sigma : [a, b] \to \mathcal{M}$ is 1/2-Hölder continuous: there exists a constant $c > 0$ such that $W_2^2(\sigma_{t_2}, \sigma_{t_1}) \leq c|t_2 - t_1|$. Together with Lemma 4.1 and the fact that $Lip(D_s) = s \leq 1$, this gives that $t \to \sigma_t^s$ is uniformly 1/2-Hölder continuous:

$$W_2^2(\sigma_{t_2}^s, \sigma_{t_1}^s) \leq W_2^2(\sigma_{t_2}, \sigma_{t_1}) \leq c|t_2 - t_1|.$$

Remark 2.11 (ii) ensures that $\{\sigma_t| \ t \in [a, b]\}$ is bounded and so there exists $\bar{c} > 0$ such that $W_2(\sigma_t, \delta_0) \leq \bar{c}$ for all $t \in [a, b]$. One can readily check that $\gamma := \left(D_{s_1} \times D_{s_2}\right)_{\#} \sigma_t \in \Gamma(\sigma_t^{s_1}, \sigma_t^{s_2})$, so

$$W_2^2(\sigma_t^{s_1}, \sigma_t^{s_2}) \leq \int_{\mathbb{R}^D \times \mathbb{R}^D} |x - y|^2 d\gamma = \int_{\mathbb{R}^D} |D_{s_1}x - D_{s_2}x|^2 d\sigma_t(x) = |s_2 - s_1|^2 \int_{\mathbb{R}^D} |x|^2 d\sigma_t(x)$$

$$\leq \bar{c}|s_2 - s_1|^2.$$

Thus $s \to \sigma_t^s$ is 1-Lipschitz. Consequently $(t, s) \to \sigma_t^s$ is 1/2-Hölder continuous. This, together with Lemma 5.10, yields the proof. QED.

Set

$$V(s, t) := \bar{\Lambda}_{\sigma_t^s}(v_t^s), \qquad W(s, t) := \bar{\Lambda}_{\sigma_t^s}(w_t^s).$$

The following proposition is extracted from [5] Theorem 8.3.1 and Proposition 8.4.6.

Proposition 5.27. *Let $\sigma \in AC_2(a, b; \mathcal{M})$ and let v be its velocity of minimal norm. Let $\mathcal{N}_1$ be the set of t such that v_t fails to be in $T_{\sigma_t}\mathcal{M}$. Let $\mathcal{N}_2$ be the set of $t \in [a, b]$ such that $\left(\pi^1 \times (\pi^2 - \pi^1)/h\right)_{\#} \eta_h$ fails to converge to $(Id \times v_t)_{\#}\sigma_t$ in the Wasserstein space $\mathcal{M}(\mathbb{R}^D \times \mathbb{R}^D)$, for some $\eta_h \in \Gamma_o(\sigma_t, \sigma_{t+h})$. Let $\mathcal{N}$ be the union of $\mathcal{N}_1$ and $\mathcal{N}_2$. Then $\mathcal{L}^1(\mathcal{N}) = 0$.*

We can now study the derivatives of V and W, as follows.

Lemma 5.28. *For each $t \in (a, b) \setminus \mathcal{N}$, the function $V(\cdot, t)$ is differentiable on $(r, 1)$ and its derivative is bounded by a constant $L_1(r)$ depending only on σ and r. Furthermore*

$$\partial_s V(s, t) = \int_{\mathbb{R}^D} \langle \bar{A}_{\sigma_t^s}(x), \frac{v_t^s(x)}{s} \rangle d\sigma_t^s(x) + \int_{\mathbb{R}^D} \langle B_{\sigma_t^s}(x)w_t^s(x), v_t^s(x) \rangle d\sigma_t^s(x).$$

Proof: Let $C_\sigma(r)$ be as in Lemma 5.26 and let $\bar{C}_\sigma$ be as in Equation 5.32. We use Equations 5.12, 5.33 and then Hölder's inequality to obtain

$$(5.36) \qquad |V(s + h, t) - V(s, t)| \leq \frac{h}{s}||\bar{A}_{\sigma_t^s}||_{\sigma_t^s} ||v_t^s||_{\sigma_t^s} + 2c(\bar{\Lambda})W_2(\sigma_t^s, \sigma_t^{s+h})||v_t^{s+h}||_{\sigma_t^{s+h}}.$$

We combine Equations 5.34, 5.35 and 5.36 to conclude that

$$(5.37) \qquad |V(s + h, t) - V(s, t)| \leq hC_\sigma(r)||\sigma'||_\infty + 2hc(\bar{\Lambda})\bar{C}_\sigma(s + h)||\sigma'||_\infty.$$

This proves that $V(\cdot, t)$ is Lipschitz on $(r, 1)$ and that its derivative is bounded by a constant $L_1(r)$. As in Remark 5.8,

$$\lim_{h \to 0} \frac{V(s + h, t) - V(s, t)}{h} = \lim_{h \to 0} \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_{\sigma_t^s}(x), \frac{v_t^{s+h}(y) - v_t^s(x)}{h} \rangle d\gamma^h$$

$$+ \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle B_{\sigma_t^s}(x)\frac{y - x}{h}, v_t^{s+h}(y) \rangle d\gamma^h$$

$$(5.38) \qquad + \frac{1}{h} \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_{\sigma_t^{s+h}}(y) - \bar{A}_{\sigma_t^s}(x) - B_{\sigma_t^s}(x)(y - x), v_t^{s+h}(y) \rangle d\gamma^h,$$

where $\gamma_h \in \Gamma_o(\sigma_t^{s+h}, \sigma_t^s)$. By Equation 5.10, the last inequality in Equation 5.34 and Equation 5.35 we have

$$(5.39) \qquad \lim_{h \to 0} \frac{1}{h} \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_{\sigma_t^{s+h}}(y) - \bar{A}_{\sigma_t^s}(x) - B_{\sigma_t^s}(x)(y - x), v_t^{s+h}(y) \rangle \gamma^h(x, y) = 0.$$

We use Equations 5.33, 5.38, 5.39 and the fact that, for γ^h-almost every $(x, y) \in \mathbb{R}^D \times \mathbb{R}^D$, $y = (1 + h/s)x$ to conclude that

$$\lim_{h \to 0} \frac{V(s + h, t) - V(s, t)}{h} = \int_{\mathbb{R}^D} \langle \bar{A}_{\sigma_t^s}(x), \frac{v_t^s(x)}{s} \rangle d\sigma_t^s + \lim_{h \to 0} \int_{\mathbb{R}^D} \langle B_{\sigma_t^s}(x) \frac{x}{s}, (1 + \frac{h}{s}) v_t^s(x) \rangle d\sigma_t^s$$

$$= \int_{\mathbb{R}^D} \langle \bar{A}_{\sigma_t^s}(x), \frac{v_t^s(x)}{s} \rangle d\sigma_t^s + \int_{\mathbb{R}^D} \langle B_{\sigma_t^s}(x) w_t^s(x), v_t^s(x) \rangle d\sigma_t^s.$$

This proves the lemma. QED.

Lemma 5.29. *For each $s \in [r, 1]$ and $t \in (a, b) \setminus \mathcal{N}$, the function $W(s, \cdot)$ is differentiable at t and its derivative is bounded by a constant $L_2(r)$ depending only on σ and r. Furthermore,*

$$\partial_t W(s, t) = \int_{\mathbb{R}^D} \langle \bar{A}_{\sigma_t^s}(x), \frac{v_t^s(x)}{s} \rangle d\sigma_t^s(x) + \int_{\mathbb{R}^D} \langle w_t^s(x), B_{\sigma_t^s}(x) v_t^s(x) \rangle d\sigma_t^s(x).$$

Proof: As in the proof of Lemma 5.30, we have

$$|W(s, t + h) - W(s, t)| \leq \frac{||\bar{A}_{\sigma_t^s}||_{\sigma_t^s}}{s} W_2(\sigma_{t+h}^s, \sigma_t^s) + 2c(\bar{\Lambda}) W_2(\sigma_{t+h}^s, \sigma_t^s) ||w_{t+h}^s||_{\sigma_{t+h}^s}.$$

This gives

$$|W(s, t + h) - W(s, t)| \leq h(1 + 2c(\bar{\Lambda})) \bar{C}_\sigma,$$

proving that $W(s, \cdot)$ is Lipschitz on (a, b) and that its derivative is bounded by a constant $L_2(r)$. For fixed $s \in [r, 1]$ and $t \in (a, b) \setminus \mathcal{N}$, let $\gamma_h^s \in \Gamma_o(\sigma_t^s, \sigma_{t+h}^s)$. Then

$$\lim_{h \to 0} \frac{W(s, t + h) - W(s, t)}{h}$$

$$= \lim_{h \to 0} \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_{\sigma_t^s}(x), \frac{w_{t+h}^s(y) - w_t^s(x)}{h} \rangle d\gamma_h^s(x, y)$$

$$+ \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle B_{\sigma_t^s}(x) \frac{y - x}{h}, w_{t+h}^s(y) \rangle d\gamma_h^s(x, y)$$

$$+ \frac{1}{h} \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_{\sigma_{t+h}^s}(y) - \bar{A}_{\sigma_t^s}(x) - B_{\sigma_t^s}(x)(y - x), w_{t+h}^s(y) \rangle \gamma_h^s(x, y).$$

However,

$$\lim_{h \to 0} \frac{1}{h} \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_{\sigma_t^s}(y) - \bar{A}_{\sigma_t^s}(x) - B_{\sigma_t^s}(x)(y - x), w_{t+h}^s(y) \rangle \gamma_h^s(x, y) = 0.$$

We then use the fact that $w_t^s(z) = z/s$ to get

$$(5.40) \qquad \lim_{h \to 0} \frac{W(s, t + h) - W(s, t)}{h} = \lim_{h \to 0} \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_{\sigma_t^s}(x), \frac{y - x}{sh} \rangle d\gamma_h^s(x, y)$$

$$+ \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle B_{\sigma_t^s}(x) \frac{y - x}{h}, \frac{y}{s} \rangle d\gamma_h^s(x, y).$$

To conclude the lemma it suffices to show that if $t \in (a,b) \setminus \mathcal{N}$ and $\gamma_h^s \in \Gamma_o(\sigma_t^s, \sigma_{t+h}^s)$ then $\left(\pi^1 \times (\pi^2 - \pi^1)/h\right)_{\#} \gamma_h^s$ converges to $(Id \times v_t^s)_{\sigma_t^s}$ in $\mathcal{M}(\mathbb{R}^D \times \mathbb{R}^D)$ as h tends to 0. Set

$$\gamma_h := \left(D_s^{-1} \times D_s^{-1}\right)_{\#} \gamma_h^s.$$

Since

$$\pi^1 \circ \left(D_s^{-1} \times D_s^{-1}\right) = D_s^{-1} \circ \pi^1 \quad \text{and} \quad \pi^2 \circ \left(D_s^{-1} \times D_s^{-1}\right) = D_s^{-1} \circ \pi^2,$$

we conclude that $\gamma_h \in \Gamma(\sigma_t, \sigma_{t+h})$. By the fact that the support of γ_h^s is cyclically monotone we have that the support of γ_h is also cyclically monotone. Hence $\gamma_h \in \Gamma_o(\sigma_t, \sigma_{t+h})$. We have

$$\left(\pi^1 \times \frac{\pi^2 - \pi^1}{h}\right)_{\#} \gamma_h^s = \left(D_s \times D_s\right)_{\#}\left((\pi^1 \times \frac{\pi^2 - \pi^1}{h})_{\#} \gamma_h\right) \to \left(D_s \times D_s\right) \circ (Id \times v_t)_{\#}\sigma_t$$

$$= (Id \times v_t^s)_{\#}\sigma_t^s.$$

This, together with Equation 5.40, gives

$$\lim_{h \to 0} \frac{W(s, t+h) - W(s,t)}{h}$$

$$= \int_{\mathbb{R}^D} \langle \bar{A}_{\sigma_t^s}(x), \frac{v_t^s(x)}{s}\rangle d\sigma_t^s(x) + \int_{\mathbb{R}^D} \langle B_{\sigma_t^s}(x) v_t^s(x), \frac{x}{s}\rangle d\sigma_t^s(x)$$

$$= \int_{\mathbb{R}^D} \langle \bar{A}_{\sigma_t^s}(x), \frac{v_t^s(x)}{s}\rangle d\sigma_t^s(x) + \int_{\mathbb{R}^D} \langle B_{\sigma_t^s}(x) v_t^s(x), w_t^s(x)\rangle d\sigma_t^s(x).$$

QED.

Corollary 5.30. *For each $s \in (r, 1)$ and $t \in (a,b) \setminus \mathcal{N}$ we have*

$$\partial_t\left(\bar{\Lambda}_{\sigma_t^s}(w_t^s)\right) - \partial_s\left(\bar{\Lambda}_{\sigma_t^s}(v_t^s)\right) = d\bar{\Lambda}_{\sigma_t^s}(v_t^s, w_t^s).$$

Proof: This corollary is a direct consequence of Lemmas 5.12, 5.28 and 5.29.   QED.

Remark 5.31. Proposition 5.2 was a direct consequence of Equations 5.3 and 5.4. Those equations depended strongly on the smoothness of v_t^s and w_t^s in all variables. In this section we have removed all smoothness assumptions on v_t^s. Specifically, now we do not know that ∇v_t^s nor $\partial_t v_t^s$ exist. However, Equation 5.33 ensures that v_t^s is differentiable with respect to s, in particular establishing the inequality $||v_t^{s+h} \circ \pi^2 - v_t^s \circ \pi^1||_{\gamma^h} \leq h||\sigma'||_\infty$. This inequality is crucial for the proof of Lemma 5.28.

Lemma 5.28 is the analogue of Equation 5.4. Notice that the first term (respectively, the second term) on the right hand side of Lemma 5.28 corresponds to the second term (respectively, the first term) on the right hand side of Equation 5.4. Likewise, Lemma 5.29 is the analogue of Equation 5.3.

Theorem 5.32 (Green's formula on the annulus). *Consider in $\mathcal{M}$ the annulus $S(s,t) = D_{s\#}\sigma_t$ for $(s,t) \in [r,1] \times [0,T]$. Let ∂S denote its boundary, defined as the union of the negatively oriented curves $S(r, \cdot)$, $S(\cdot, T)$ and the positively oriented curves $S(1, \cdot)$, $S(\cdot, 0)$. Then*

$$\int_S d\bar{\Lambda} = \int_{\partial S} \bar{\Lambda}.$$

Proof: We use Corollary 5.30 to obtain

$$\int_S d\bar\Lambda = \int_0^T dt \int_r^1 d\bar\Lambda_{S(s,t)}(v_t^s, w_t^s) ds = \int_0^T dt \int_r^1 \left[\partial_t\left(\bar\Lambda_{S(s,t)}(w_t^s)\right) - \partial_s\left(\bar\Lambda_{S(s,t)}(v_t^s)\right)\right] ds$$

$$(5.41) \quad = \int_r^1 \left(\bar\Lambda_{S(s,T)}(w_T^s) - \bar\Lambda_{S(s,0)}(w_0^s)\right) ds - \int_0^T \left(\bar\Lambda_{S(1,t)}(v_t^1) - \bar\Lambda_{S(r,t)}(v_t^r)\right) dt = \int_{\partial S} \bar\Lambda.$$

<div align="right">QED.</div>

Corollary 5.33. *If we further assume that $\bar\Lambda$ is a closed pseudo 1-form and that $\sigma_0 = \sigma_T$, then $\int_\sigma \bar\Lambda = 0$.*

Proof: For $s \in [r, 1]$ define

$$l(s) = \int_0^T \bar\Lambda_{S(s,t)}(v_t^s) dt, \qquad \bar l(t) = \int_r^1 \bar\Lambda_{S(s,t)}(w_t^s) ds.$$

Since $w_T^s = w_0^s$ and $\sigma_T^s = D_{s\#}\sigma_T = D_{s\#}\sigma_0 = \sigma_0^s$, we have $\bar l(T) = \bar l(0)$. This, together with Equation 5.41 and the fact that $d\bar\Lambda = 0$, yields $\int_0^T \bar\Lambda_{\sigma_t}(v_t) dt = l(1) = l(r)$. But
(5.42)
$$|l(r)| \leq \int_0^T |\bar\Lambda_{S(s,t)}(v_t^r)| dt \leq \int_0^T ||\bar A_{S(s,t)}||_{S(s,t)} ||v_t^r||_{S(s,t)} dt \leq r||\sigma'||_\infty \int_0^T ||\bar A_{S(s,t)}||_{S(s,t)} dt,$$

where we have used the last inequality in Equation 5.34. The first inequality in Equation 5.32 shows that, for r small enough, $\{S(s,t)\}_{t\in\times[0,T])}$ is contained in a small ball centered at δ_0. But Lemma 5.10 gives that $\mu \to ||\bar A_\mu||_\mu$ is continuous at δ_0. Thus there exist constants c and r_0 such that $||\bar A_{S(s,t)}||_{S(s,t)} \leq c$ for all $t \in [0,T]$ and all $r < r_0$. We can now exploit Equation 5.42 to obtain

$$|l(1)| = \liminf_{r\to 0} |l(r)| \leq \liminf_{r\to 0} rTc||\sigma'||_\infty = 0.$$

<div align="right">QED.</div>

Corollary 5.34. *Let $\bar\Lambda$ be a regular pseudo 1-form on $\mathcal{M}$. Let Λ denote the corresponding 1-form on $\mathcal{M}$, defined by restriction. Assume $\bar\Lambda$ is closed, i.e. $d\bar\Lambda = 0$. Then Λ is exact, i.e. there exists a differentiable function F on $\mathcal{M}$ such that $dF = \Lambda$.*

Proof: Fix $\mu \in \mathcal{M}$. Let σ be any curve in $AC_2(a, b; \mathcal{M})$ such that $\sigma_a = \delta_0$ and $\sigma_b = \mu$. Assume that v is its velocity of minimal norm and that $\sup_{(a,b)} ||v_t||_{\sigma_t} < \infty$. By Corollary 5.33, $\int_\sigma \bar\Lambda$ depends only on μ, i.e. it is independent of the path σ. Also, Remark 5.24 ensures that $\int_\sigma \bar\Lambda$ is independent of a, b. It is thus meaningful to define

$$F(\mu) := \int_\sigma \bar\Lambda.$$

We now want to show that F is differentiable. Fix $\mu, \nu \in \mathcal{M}$ and $\gamma \in \Gamma_o(\mu, \nu)$. Define $\sigma_t := ((1-t)\pi^1 + t\pi^2)_{\#}\gamma$. Then $\sigma : [0, 1] \to \mathcal{M}$ is a constant speed geodesic between μ and ν. Let v_t denote its velocity of minimal norm. Clearly,

$$(5.43) \qquad\qquad F(\nu) - F(\mu) = \int_0^1 \bar\Lambda_{\sigma_t}(v_t) dt.$$

Let $\bar{\gamma} : \mathbb{R}^D \to \mathbb{R}^D$ denote the *barycentric projection* of γ, cf. [5] Definition 5.4.2. Set $v := \bar{\gamma} - Id$. Then $\gamma_t := (\pi^1, (1-t)\pi^1 + t\pi^2)_{\#}\gamma \in \Gamma_o(\sigma_0, \sigma_t)$ and

$$\bar{\Lambda}_{\sigma_t}(v_t) - \bar{\Lambda}_{\sigma_0}(v) = \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_{\sigma_0}(x), v_t(y) - v(x) \rangle + \langle B_{\sigma(0)}(x)(y-x), v_t(y) \rangle d\gamma_t(x,y)$$

$$+ \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_{\sigma_t}(y) - \bar{A}_{\sigma_0}(x) - B_{\sigma_0}(x)(y-x), v_t(y) \rangle d\gamma_t(x,y).$$

By Equation (5.7) and Hölder's inequality,

$$\left| \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_{\sigma_t}(y) - \bar{A}_{\sigma_0}(x) - B_{\sigma_0}(x)(y-x), v_t(y) \rangle d\gamma_t(x,y) \right| \leq o(W_2(\sigma_0, \sigma_t)) \, ||v_t||_{\sigma_t}.$$

It is well known (cf. [5] Lemma 7.2.1) that if $0 < t \leq 1$ then there exists a unique optimal transport map T_t^1 between σ_t and σ_1, *i.e.* $\Gamma_o(\sigma_t, \sigma_1) = \{(Id \times T_t^1)_{\#}\sigma_t\}$. One can check that $v_t(y) = \frac{T_t^1(y)-y}{1-t}$ and $||v_t||_{\sigma_t} = W_2(\sigma_t, \sigma_1)/(1-t) = W_2(\sigma_0, \sigma_1)$. Thus

$$\int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_{\sigma_0}(x), v_t(y) - v(x) \rangle d\gamma_t(x,y) = \int \langle \bar{A}_{\sigma_0}(x), \frac{T_t^1(y)-y}{1-t} - (\bar{\gamma}(x) - x) \rangle d\gamma_t(x,y)$$

$$= \int \langle \bar{A}_{\sigma_0}(x), \frac{z - ((1-t)x + tz)}{1-t} - (z-x) \rangle d\gamma(x,z)$$

$$= 0.$$

Similarly,

$$\int_{\mathbb{R}^D \times \mathbb{R}^D} \langle B_{\sigma_0}(x)(y-x), v_t(y) \rangle d\gamma_t(x,y) = t \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle B_{\sigma_0}(x)(z-x), z-x \rangle d\gamma(x,y)$$

$$= o(W_2(\sigma_0, \sigma_1)) = o(W_2(\mu, \nu)).$$

Combining these equations shows that

$$(5.44) \qquad\qquad \bar{\Lambda}_{\sigma_t}(v_t) - \bar{\Lambda}_{\sigma_0}(v) = o(W_2(\mu, \nu)).$$

Notice that (5.44) is independent of t. Combining (5.43) and (5.44) we find

$$F(\nu) = F(\mu) + \bar{\Lambda}_{\sigma_0}(v) + \int_0^1 \bar{\Lambda}_{\sigma_t}(v_t) - \bar{\Lambda}_{\sigma_0}(v) dt$$

$$= F(\mu) + \bar{\Lambda}_{\sigma_0}(v) + o(W_2(\mu, \nu))$$

$$= F(\mu) + \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_{\sigma_0}(x), y - x \rangle d\gamma(x,y) + o(W_2(\mu, \nu)).$$

As in Definition 4.9, this proves that F is differentiable and that $\nabla_\mu F = \pi_\mu(\bar{A}_\mu)$. Thus $dF = \Lambda$.   QED.

Example 5.35. Assume $\bar{\Lambda}$ is a linear pseudo 1-form, *i.e.* $\bar{\Lambda}(\cdot) = \int_{\mathbb{R}^D} \langle \bar{A}, \cdot \rangle d\mu$ for some $\bar{A} \in \mathcal{X}_c$. According to Example 4.15, if $d\bar{\Lambda} = 0$ then $\int_{\mathbb{R}^D} \langle (\nabla \bar{A} - \nabla \bar{A}^T) \cdot, \cdot \rangle d\mu = 0$ on $T_\mu \mathcal{M}$. Restricting to Dirac measures proves that $\nabla \bar{A}$ is symmetric so $\bar{A}$ is a gradient vector field. In other words, any closed linear pseudo 1-form is actually a linear 1-form.

5.5. **Example: 1-forms on the space of discrete measures.** Fix an integer $n \geq 1$. Given $x_1, \cdots, x_n \in \mathbb{R}^D$, set $\mathbf{x} := (x_1, \cdots, x_n)$ and $\mu_{\mathbf{x}} := 1/n \sum_{i=1}^n \delta_{x_i}$. Let M denote the set of such measures and TM denote its tangent bundle, cf. Examples 2.2 and 2.8. Choose a regular pseudo 1-form $\bar{\Lambda}$ on $\mathcal{M}$. By restriction we obtain a 1-form α on M, defined by $\alpha_{\mathbf{x}} := \bar{\Lambda}_{\mu_{\mathbf{x}}}$. Let $A : \mathbb{R}^{nD} \to \mathbb{R}^{nD}$ be defined by

$$A(\mathbf{x}) = (A_1(\mathbf{x}), \cdots, A_n(\mathbf{x})) := \left(\bar{A}_{\mu_{\mathbf{x}}}(x_1), \cdots, \bar{A}_{\mu_{\mathbf{x}}}(x_n) \right).$$

Notice that if $X = (X_1, \cdots, X_n) \in \mathbb{R}^{nD}$ satisfies $X_i = X_j$ whenever $x_i = x_j$ then $\alpha_{\mathbf{x}}(X) = \frac{1}{n}\langle A(\mathbf{x}), X \rangle$. Now define a $nD \times nD$ matrix $B(\mathbf{x})$ by setting

$$(5.45) \qquad B_{k+i,k+j} := \left(B_{\mu_{\mathbf{x}}}(x_{k+1}) \right)_{ij}, \qquad \text{for } k = 0, \cdots n-1, \quad i,j = 1, \cdots, D,$$

$$(5.46) \qquad B_{l,m} := 0 \quad \text{if} \quad (l,m) \notin \{(k+i, k+j) : k = 0, \cdots n-1, \quad i,j = 1, \cdots, D\}.$$

Proposition 5.36. *The map $A : \mathbb{R}^{nD} \to \mathbb{R}^{nD}$ is differentiable and $\nabla A(\mathbf{x}) = B(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^{nD}$.*

Proof: Let $\mathbf{x} = (x_1, \cdots, x_n) \in \mathbb{R}^{nD}$. Set $r := \min_{x_i \neq x_j} |x_i - x_j|$. If $\mathbf{y} = (y_1, \cdots, y_n) \in \mathbb{R}^{nD}$ and $|\mathbf{y} - \mathbf{x}| < r/2$ then $\Gamma_o(\mu_{\mathbf{x}}, \mu_{\mathbf{y}})$ has a single element $\gamma_{\mathbf{y}} = 1/n \sum_{i=1}^n \delta_{(x_i,y_i)}$ and $nW_2^2(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}) = |\mathbf{y} - \mathbf{x}|^2$. By Equation 5.7,

$$(5.47) \qquad |A(\mathbf{y}) - A(\mathbf{x}) - B(\mathbf{x})(\mathbf{y} - \mathbf{x})|^2 = n \, o(\frac{|\mathbf{y} - \mathbf{x}|^2}{n}).$$

This concludes the proof. QED.

Lemma 5.37. *Suppose $\mathbf{x} = (x_1, \cdots, x_n) \in \mathbb{R}^{nD}$ and $X = (X_1, \cdots, X_n)$, $Y = (Y_1, \cdots, Y_n) \in \mathbb{R}^{nD}$ are such that $X_i = X_j$, $Y_i = Y_j$ whenever $x_i = x_j$. Then*

$$d\bar{\Lambda}_{\mu_{\mathbf{x}}}(X, Y) = d\alpha_{\mathbf{x}}(X, Y).$$

Proof: We use Lemma 5.15 and Equations 5.45, 5.46 to obtain

$$d\bar{\Lambda}_{\mu_{\mathbf{x}}}(X, Y) = \sum_{k=1}^n \left\langle (B_{\mu_{\mathbf{x}}}(x_k) - B_{\mu_{\mathbf{x}}}(x_k)^T)X_k, Y_k \right\rangle = d\alpha_{\mathbf{x}}(X, Y).$$

QED.

Corollary 5.38. *Suppose that $\mathbf{r} = (r_1, \cdots, r_n) \in C^2([0,T], \mathbb{R}^{nD})$ and set $\sigma_t := 1/n \sum_{i=1}^n \delta_{r_i(t)}$. If $\bar{\Lambda}$ is closed and $\sigma_0 = \sigma_T$ then $\int_\sigma \alpha = 0$.*

Proof: This is a direct consequence of Corollary 5.33. QED.

Remark 5.39. One can check by direct computation that, for a surface $\mathbf{x} = \mathbf{x}(s,t)$ in M, the familiar identity $\partial_t(\alpha_{\mathbf{x}}(\partial_s\mathbf{x})) - \partial_s(\alpha_{\mathbf{x}}(\partial_t\mathbf{x})) = d\alpha_{\mathbf{x}}(\partial_t\mathbf{x}, \partial_s\mathbf{x})$ holds. Together with Lemma 5.37, this is the analogue of Corollary 5.30 which we used to prove Theorem 5.32.

Remark 5.40. Notice that the assumption $\sigma_0 = \sigma_T$ is weaker than $\mathbf{r}(0) = \mathbf{r}(T)$.

5.6. **Discussion.** As mentioned in Section 2.2, the space $\mathcal{M}$ is convex. This is true not only for probability measures on $\mathbb{R}^D$ but also for probability measures on any manifold M. It is thus trivially true that all cohomology groups $H^k(\mathcal{M}; \mathbb{R})$ vanish if these groups are defined purely in terms of the topology on $\mathcal{M}$.

In this paper, however, we are concerned with the differentiable structure on $\mathcal{M}$. As seen in Section 3.2, from this point of view $\mathcal{M}$ is the union of smooth manifolds $\mathcal{O}$, defined as the orbits of the $\mathrm{Diff}_c(\mathbb{R}^D)$-action on $\mathcal{M}$. Given any such orbit $\mathcal{O}$, it is then interesting to define and calculate the first cohomology group $H^1(\mathcal{O}; \mathbb{R})$. Notice that these orbits are in general not convex, so the above reasoning does not apply. In the case $M = \mathbb{R}^D$, however, Corollary 5.34 can intuitively be interpreted as a vanishing result for $H^1(\mathcal{O}; \mathbb{R})$ if we think of cohomology in the differentiable sense, *i.e.* in the sense of de Rham, as follows.

Let Λ be a regular 1-form on $\mathcal{M}$. Recall from Definition 4.12 that Λ is closed iff $\bar{\Lambda}$ is closed. Then Corollary 5.34 shows that if Λ is closed then it is exact, *i.e.* $\Lambda = dF$ for some differentiable $F : \mathcal{M} \to \mathbb{R}$. Now choose an orbit $\mathcal{O}$. Given any $\mu \in \mathcal{O}$, recall from Section 3.2 that $T_\mu \mathcal{M}$ can be thought of as the tangent space to $\mathcal{O}$ at the point μ. In this sense the restriction of F to $\mathcal{O}$ is still differentiable and $d(F_{|\mathcal{O}}) = (dF)_{|\mathcal{O}}$ so the restriction of Λ to $\mathcal{O}$ is also exact. Roughly speaking, Corollary 5.34 thus shows that the first *de Rham cohomology* group $H^1(\mathcal{O}; \mathbb{R})$ of each orbit vanishes. It may be useful to point out that if μ is a Dirac measure then $\mathcal{O}_\mu = \mathbb{R}^D$, so at least in this case the above vanishing result makes sense.

It is reasonable to expect that most of the theory of [5] can be extended to probability measure spaces on other manifolds M. In this case, many of the results of this paper should also extend. However the above example, where $\mathcal{O}$ is the space of Dirac measures on M, shows that one should not expect $H^1(\mathcal{O}; \mathbb{R})$ to vanish in general. In this sense our results are specific to the choice $M = \mathbb{R}^D$. Another way to see this is as follows. The proof of Corollary 5.34 relied on the construction of certain "annuli" built using maps D_s of $\mathbb{R}^D$. These maps exist only because $\mathbb{R}^D$ is contractible. Such a construction would not be possible on other manifolds.

The following considerations also support the above interpretation of Corollary 5.34. Recall that, for a finite-dimensional manifold M, the first de Rham cohomology group is closely related to the topology of M, as follows: $H^1(M; \mathbb{R}) = \mathrm{Hom}(\pi_1(M), \mathbb{R})$, where the latter denotes the space of group homomorphisms from the first fundamental group $\pi_1(M)$ to $\mathbb{R}$. In our case, an orbit $\mathcal{O}$ is generally not a manifold in any rigorous sense so it is not *a priori* clear that there exists any relationship between our $H^1(\mathcal{O}; \mathbb{R})$ and $\pi_1(\mathcal{O})$. However, we can formally prove the topological counterpart of our vanishing result as follows.

Let G be a finite-dimensional Lie group and H be a closed subgroup. Recall that there exists a *homotopy long exact sequence*

$$\cdots \to \pi_1(H) \to \pi_1(G) \to \pi_1(G/H) \to \pi_0(H) \to \pi_0(G) \ldots,$$

cf. *e.g.* [11], VII.5. Now assume G is connected, *i.e.* $\pi_0(G) = 1$. We can then dualize the final part of this sequence obtaining a new exact sequence

(5.48) $1 \to \mathrm{Hom}(\pi_0(H), \mathbb{R}) \to \mathrm{Hom}(\pi_1(G/H), \mathbb{R}) \to \mathrm{Hom}(\pi_1(G), \mathbb{R}).$

Now choose $\mu \in \mathcal{O}$ and set $G := \mathrm{Diff}_c(\mathbb{R}^D)$ and $H := \mathrm{Diff}_{c,\mu}(\mathbb{R}^D)$ so that $G/H \simeq \mathcal{O}$. In many cases it is known that $\pi_1(G)$ is finite: specifically, this is true at least for $D = 1, 2, 3$ and $D \geq 12$, cf. [6] for related results. Let us assume that H has a finite number of components and that the homotopy long exact sequence is still valid in this infinite-dimensional setting.

Sequence 5.48 then becomes

$$1 \to 1 \to \operatorname{Hom}(\pi_1(\mathcal{O}), \mathbb{R}) \to 1,$$

so by exactness $\operatorname{Hom}(\pi_1(\mathcal{O}), \mathbb{R})$ must also be trivial.

To conclude, it is also interesting to examine the relationship between Corollary 5.34 and *invariant cohomology*, in the sense of Section A.3. Recall from Proposition A.17 that the first cohomology group of an orbit is a subgroup of the corresponding first invariant cohomology group. The statement that the first invariant cohomology vanishes is thus stronger than the statement that the first cohomology of the orbit vanishes. Now choose any orbit $\mathcal{O}$ in $\mathcal{M}$. According to Section 4.3, the corresponding invariant cohomology should be defined in terms of regular pseudo 1-forms. To prove that the invariant cohomology vanishes would thus require showing that $d\bar{\Lambda} = 0$ implies $\bar{\Lambda} = dF$, for some differentiable function $F : \mathcal{M} \to \mathbb{R}$. Since dF is a 1-form, such a result would imply that any closed pseudo 1-form is a 1-form. Corollary 5.34 does not achieve this. On the other hand, it is not clear that such a result should even be expected.

6. A SYMPLECTIC FOLIATION OF $\mathcal{M}$

In Section 3.2 we used the action of the group of diffeomorphisms $\operatorname{Diff}_c(\mathbb{R}^D)$ to build a foliation of $\mathcal{M}$: this allowed us to formally reconstruct the differential calculus on $\mathcal{M}$. We now specialize to the case $D = 2d$. In this case the underlying manifold $\mathbb{R}^{2d}$ has a natural extra structure, the *symplectic structure* ω. The goal of this section is to use this extra data to build a second, finer, foliation of $\mathcal{M}$; we then prove that each leaf of this foliation admits a symplectic structure Ω. The foliation is obtained via a smaller group of diffeomorphisms defined by ω, the *Hamiltonian diffeomorphisms*. Section 6.1 provides an introduction to this group, cf. [36] or [32] for details.

6.1. **The group of Hamiltonian diffeomorphisms.** Recall that a *symplectic structure* on a (possibly infinite-dimensional) vector space V is a 2-form $\omega : V \times V \to \mathbb{R}$ such that

$$(6.1) \qquad\qquad \omega^\flat : V \to V^*, \quad v \mapsto i_v \omega := \omega(v, \cdot)$$

is injective. If V is finite-dimensional then $\omega^\flat$ is an isomorphism; we will denote its inverse by $\omega^\sharp$.

Let M be a smooth manifold of dimension $D := 2d$. A *symplectic structure* on M is a smooth closed 2-form ω satisfying Equation 6.1 at each tangent space $V = T_x M$; equivalently, such that ω^d is a volume form on M. Notice that, since $d\omega = 0$, Cartan's formula A.13 shows that $\mathcal{L}_X \omega = d i_X \omega$. Throughout this section, to simplify notation, we will drop the difference between compact and noncompact manifolds but the reader should keep in mind that in the latter case we always silently restrict our attention to maps and vector fields with compact support.

Consider the set of *symplectomorphisms* of M, *i.e.*

$$\operatorname{Symp}(M) := \{\phi \in \operatorname{Diff}(M) : \phi^* \omega = \omega\}.$$

This is clearly a subgroup of $\operatorname{Diff}(M)$. Using the methods of Section A.4 (see in particular Remark A.21) one can show that it has a Lie group structure. Its tangent space at Id, thus its Lie algebra, is by definition isomorphic to the space of closed 1-forms on M. Via $\omega^\sharp$

and Formula A.13 this space is isomorphic to the space of *symplectic* or *locally Hamiltonian* vector fields, *i.e.*

$$\operatorname{Symp} \mathcal{X} := \{X \in \mathcal{X}(M) : \mathcal{L}_X \omega = 0\}.$$

Remark 6.1. Equation A.9 confirms that $\operatorname{Symp} \mathcal{X}$ is closed under the bracket operation, *i.e.* that it is a Lie subalgebra of $\mathcal{X}(M)$. Equation A.10 confirms that $\operatorname{Symp} \mathcal{X}$ is closed under the push-forward operation, *i.e.* under the adjoint representation of $\operatorname{Symp}(M)$ on $\operatorname{Symp} \mathcal{X}$, cf. Lemma A.19.

We say that a vector field X on M is *Hamiltonian* if the corresponding 1-form $\xi := \omega(X, \cdot)$ is exact: $\xi = df$. We then write $X = X_f$. This defines the space of *Hamiltonian vector fields* $\operatorname{Ham} \mathcal{X}$. It is useful to rephrase this definition as follows. Consider the map

$$(6.2) \qquad C^\infty(M) \to \mathcal{X}(M), \quad f \mapsto df \simeq X_f := \omega^\sharp(df).$$

The Hamiltonian vector fields are the image of this map. This map is linear. It is not injective: its kernel is the space of functions constant on M. In Section 7.1 we will start referring to these functions as the *Casimir functions* for the map of Equation 6.2.

Remark 6.2. We can rephrase the properties of the map of Equation 6.2 by saying that there exists a *short exact sequence*

$$(6.3) \qquad 0 \to \mathbb{R} \to C^\infty(M) \to \operatorname{Ham} \mathcal{X} \to 0.$$

As already mentioned, the function corresponding to a given Hamiltonian vector field is well-defined only up to a constant. In some cases we can fix this constant via a *normalization, i.e.* we can build an inverse map $\operatorname{Ham} \mathcal{X} \to C^\infty(M)$. We then obtain an isomorphism between $\operatorname{Ham} \mathcal{X}$ and the space of normalized functions. For example, if M is compact we can fix this constant by requiring that f have integral zero, $\int_M f \omega^d = 0$. If instead $M = \mathbb{R}^{2d}$ and we restrict our attention as usual to Hamiltonian diffeomorphisms with compact support, we should restrict Equation 6.2 to the space $\mathbb{R} \oplus C_c^\infty(\mathbb{R}^{2d})$ of functions which are constant outside of a compact set; by restriction we then get an isomorphism $C_c^\infty(\mathbb{R}^{2d}) \simeq \operatorname{Ham} \mathcal{X}_c$.

More generally, we say that a time-dependent vector field X_t is *Hamiltonian* if $\omega(X_t, \cdot) = df_t$ for some curve of smooth functions f_t. We then say that the diffeomorphism $\phi \in \operatorname{Diff}(M)$ is *Hamiltonian* if it can be obtained as the time $t = 1$ flow of a time-dependent Hamiltonian vector field X_{f_t}, *i.e.* if $\phi = \phi_1$ and ϕ_t solves Equation A.8.

Let $\operatorname{Ham}(M)$ denote the set of Hamiltonian diffeomorphisms. It follows from Lemma A.3 that all such maps are symplectomorphisms. It is not immediately obvious that $\operatorname{Ham}(M)$ is closed under composition but it is not hard to prove that this is indeed true, cf. [36] Proposition 10.2 and Exercise 10.3. Once again, the methods of Section A.4 and Remark A.21 show that $\operatorname{Ham}(M)$ has a Lie group structure. Its tangent space at Id, thus its Lie algebra, is isomorphic to the space of exact 1-forms, which via $\omega^\sharp$ corresponds to the space of Hamiltonian vector fields.

It is a fundamental fact of Symplectic Geometry that ω defines a Lie bracket on $C^\infty(M)$, as follows:

$$\{f, g\} := \omega(X_f, X_g) = df(X_g) = \mathcal{L}_{X_g} f.$$

This operation is clearly bilinear and anti-symmetric. The fact that it satisfies the Jacobi identity, cf. Definition A.2, follows from the following standard result.

Lemma 6.3. *Let $\phi \in Symp(M)$. Then $\phi^* X_f = X_{\phi^* f}$ and $\phi^* \{f, g\} = \{\phi^* f, \phi^* g\}$. Applying this to $\phi_t \in Symp(M)$ and differentiating, it implies:*

(6.4) $$\mathcal{L}_{X_h} \{f, g\} = \{\mathcal{L}_{X_h} f, g\} + \{f, \mathcal{L}_{X_h} g\}.$$

Lemma 6.4. *The map $f \mapsto X_f$ has the following property:*

$$X_{\{f,g\}} = -[X_f, X_g].$$

Proof: It is enough to check that $dh(X_{\{f,g\}}) = -dh([X_f, X_g])$, for all $h \in C^\infty(M)$. As usual, it will simplify the notation to set $X(f) := df(X)$. In particular $X_f(h) = \{h, f\}$ and $dh([X, Y]) = X(Y(h)) - Y(X(h))$. Then:

$$\begin{aligned} X_{\{f,g\}}(h) &= \{h, \{f, g\}\} = -\{f, \{g, h\}\} - \{g, \{h, f\}\} \\ &= -\{\{h, g\}, f\} + \{\{h, f\}, g\} = -X_f(X_g(h)) + X_g(X_f(h)) \\ &= -[X_f, X_g](h). \end{aligned}$$

QED.

Recall from Section A.4 the negative sign appearing in the Lie bracket $[\cdot, \cdot]_{\mathfrak{g}}$ on vector fields. It follows from Lemma 6.4 that the map of Equation 6.2 is a Lie algebra homomorphism between $C^\infty(M)$ and the space of Hamiltonian vector fields, endowed with that Lie bracket.

Remark 6.5. Lemma 6.4 confirms that $\text{Ham}\,\mathcal{X}$ is a Lie subalgebra of $\mathcal{X}(M)$. Lemma 6.3 confirms that it is closed under symplectic push-forward, so in particular it is closed under the adjoint representation of $\text{Ham}(M)$.

Remark 6.6. Notice that $\text{Ham}(M)$ is connected by definition. If M satisfies $H^1(M, \mathbb{R}) = 0$, *i.e.* every closed 1-form is exact, then every symplectic vector field is Hamiltonian. Now assume that $\phi \in \text{Symp}(M)$ is such that there exists $\phi_t \in \text{Symp}(M)$ with $\phi_0 = Id$ and $\phi_1 = \phi$. It then follows from Lemma A.3 that ϕ is Hamiltonian, *i.e.* that the connected component of $\text{Symp}(M)$ containing the identity coincides with $\text{Ham}(M)$. In particular this applies to $M = \mathbb{R}^{2d}$, so in later sections we could just as well choose to work with (the connected component containing Id of) $\text{Symp}_c(\mathbb{R}^{2d})$ rather than with $\text{Ham}_c(\mathbb{R}^{2d})$. We choose however not to do this, so as to emphasize the fact that for general M the two groups are indeed different and that generalizing our constructions would require working with $\text{Ham}(M)$ rather than with $\text{Symp}(M)$.

Remark 6.7. In many cases it is known that $\text{Symp}(M)$ is closed in $\text{Diff}(M)$ and that $\text{Ham}(M)$ is closed in $\text{Symp}(M)$, see [36] and [39] for details.

6.2. **A symplectic foliation of $\mathcal{M}$.** The manifold $\mathbb{R}^{2d}$ has a natural symplectic structure defined by $\omega := dx^i \wedge dy^i$. Let J denote the natural *complex structure* on $\mathbb{R}^{2d}$, defined with respect to the basis $\partial x^1, \ldots, \partial x^d, \partial y^1, \ldots, \partial y^d$ by the matrix

$$J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}.$$

Notice that $\omega(\cdot, \cdot) = g(J \cdot, \cdot)$. It follows from this that Hamiltonian vector fields on $\mathbb{R}^{2d}$ satisfy the identity

(6.5) $$X_f = -J \nabla f.$$

Set $\mathcal{G} := \mathrm{Ham}_c(\mathbb{R}^{2d})$, the group of compactly-supported Hamiltonian diffeomorphisms on $\mathbb{R}^{2d}$. Let $\mathrm{Ham}\,\mathcal{X}_c$ denote the corresponding Lie algebra, *i.e.* the space of compactly supported Hamiltonian vector fields on $\mathbb{R}^{2d}$. The push-forward action of $\mathrm{Diff}_c(\mathbb{R}^{2d})$ on $\mathcal{M}$ restricts to an action of $\mathcal{G}$. The corresponding orbits and stabilizers are

$$\mathcal{O}_\mu := \{\nu \in \mathcal{M} : \nu = \phi_\# \mu, \ \text{ for some } \phi \in \mathcal{G}\}, \quad \mathcal{G}_\mu := \{\phi \in \mathcal{G} : \phi_\# \mu = \mu\}.$$

Notice that this action provides a second foliation of $\mathcal{M}$, finer than the one of Section 3.2.

Example 6.8. As in Example 2.2, let a_i $(i = 1, \ldots, n)$ be a fixed collection of positive numbers such that $\sum a_i = 1$ and $x_1, \ldots, x_n \in \mathbb{R}^{2d}$ be n distinct points. Set $\mu = \sum_{i=1}^n a_i \, \delta_{x_i} \in \mathcal{M}$ and

$$\mathcal{O} = \left\{ \sum_{i=1}^n a_i \, \delta_{\bar{x}_i} \ : \ \bar{x}_1, \ldots, \bar{x}_n \in \mathbb{R}^{2d} \quad \text{are distinct} \right\}.$$

Since smooth Hamiltonian diffeomeorphisms are one-to-one maps of $\mathbb{R}^{2d}$ it is clear that $\mathcal{O}_\mu \subseteq \mathcal{O}$. Given any $\bar{x}_1 \in \mathbb{R}^{2d} \setminus \{x_2, \ldots, x_n\}$ one can show that there exists a Hamiltonian diffeomorphism ϕ with compact support such that $\phi(x_1) = \bar{x}_1$ and $\phi(x_i) = x_i$ for $i \neq 1$. Thus, setting $\bar{\mu} := a_1 \, \delta_{\bar{x}_1} + \sum_{i=2}^n a_i \, \delta_{x_i}$, we see that $\bar{\mu} \in \mathcal{O}_\mu$. Repeating the argument $n - 1$ times we conclude that $\mathcal{O} \subseteq \mathcal{O}_\mu$, so $\mathcal{O} = \mathcal{O}_\mu$.

Definition 6.9. Let $\mu \in \mathcal{M}$. Consider the $L^2(\mu)$-closure $\overline{\mathrm{Ham}\,\mathcal{X}_c}^\mu$ of $\mathrm{Ham}\,\mathcal{X}_c$. We can restrict the operator div_μ to this space; we will continue to denote its kernel $\mathrm{Ker}(div_\mu)$. We define the *symplectic tangent subspace* at μ to be the space

$$T_\mu \mathcal{O} := \overline{\mathrm{Ham}\,\mathcal{X}_c}^\mu / \mathrm{Ker}(div_\mu) \subseteq L^2(\mu) / \mathrm{Ker}(div_\mu).$$

Recall from Remark 2.7 the identification $\pi_\mu : L^2(\mu) / \mathrm{Ker}(div_\mu) \to T_\mu \mathcal{M}$. By restriction this allows us to identify $T_\mu \mathcal{O}$ with the subspace $\pi_\mu(\overline{\mathrm{Ham}\,\mathcal{X}_c}^\mu) \subseteq T_\mu \mathcal{M}$. We define the *pseudo symplectic distribution* on $\mathcal{M}$ to be the union of all spaces $\overline{\mathrm{Ham}\,\mathcal{X}_c}^\mu$, for $\mu \in \mathcal{M}$. It is a subbundle of $\mathcal{T}\mathcal{M}$. We define the the *symplectic distribution* on $\mathcal{M}$ to be the union of all spaces $T_\mu \mathcal{O}$, for $\mu \in \mathcal{M}$. Up to the above identification, it is a subbundle of $T\mathcal{M}$.

Remark 6.10. Recall that in general a Hilbert space projection will not necessarily map closed subspaces to closed subspaces. Thus it is not clear that $\pi_\mu(\overline{\mathrm{Ham}\,\mathcal{X}_c}^\mu)$ is closed in $T_\mu \mathcal{M}$. On the other hand, the space $T_\mu \mathcal{O}$ of Definition 6.9 has a natural Hilbert space structure. In other words, from the Hilbert space point of view the two notions of $T_\mu \mathcal{O}$ introduced in Definition 6.9 are not necessarily equivalent. This is in contrast with the two notions of $T_\mu \mathcal{M}$, cf. Definition 2.5 and Remark 2.7.

Remark 6.11. Formally speaking the symplectic distribution is *integrable* because it is the set of tangent spaces of the smooth foliation defined by the action of $\mathcal{G}$.

Example 6.12. It is interesting to compare the space $\overline{\mathrm{Ham}\,\mathcal{X}_c}^\mu$ to the subspaces defined by Decomposition 2.5. For example, let $\mu = \delta_x$. Recall from Example 2.8 that for any $\xi \in L^2(\mu)$ there exists $\bar{\varphi} \in C_c^\infty$ such that $\xi(x) = \nabla\bar{\varphi}(x)$. Thus $\overline{\nabla C_c^\infty}^\mu = L^2(\mu)$. Now choose any $X \in L^2(\mu)$ and apply this construction to $\xi := JX$. Then $X(x) = -J\nabla\bar{\varphi}(x)$, so $\overline{\mathrm{Ham}\,\mathcal{X}_c}^\mu = L^2(\mu)$. This is the infinitesimal version of Example 6.8. In particular, $\overline{\mathrm{Ham}\,\mathcal{X}_c}^\mu = \overline{\nabla C_c^\infty}^\mu$.

The "opposite extreme" is represented by the absolutely continuous case $\mu = \rho\mathcal{L}$, for some $\rho > 0$. In this case if a Hamiltonian vector field is a gradient vector field, *e.g.* $-J\nabla v = \nabla u$,

then the function $u + iv$ is holomorphic on $\mathbb{C}^d$, so u and v are *pluriharmonic* functions on $\mathbb{R}^{2d}$ in the sense of the theory of several complex variables. This is a very strong condition: in particular, it implies that u and v are harmonic. Thus $\operatorname{Ham} \mathcal{X}_c \cap \nabla C_c^\infty = \{0\}$.

We can also compare $\overline{\operatorname{Ham} \mathcal{X}_c}^\mu$ with $\operatorname{Ker}(div_\mu)$. When $\mu = \delta_x$ we saw in Example 2.8 that $\operatorname{Ker}(div_\mu) = \{0\}$, so $\overline{\operatorname{Ham} \mathcal{X}_c}^\mu \cap \operatorname{Ker}(div_\mu) = \{0\}$. On the other hand, assume $\mu = \rho \mathcal{L}$ for some $\rho > 0$. Then $div_\mu(X) = \rho \, div(X) + \langle \nabla \rho, X \rangle$. Choose $X = -J\nabla f$. Then $div(X) = 0$ so $div_\mu(X) = 0$ iff $\langle \nabla \rho, -J\nabla f \rangle = 0$. Choosing in particular $f = \rho$ shows that $\operatorname{Ham} \mathcal{X}_c \cap \operatorname{Ker}(div_\mu) \neq \{0\}$.

We now want to show that each $T_\mu \mathcal{O}$ has a natural symplectic structure; this will justify the terminology of Definition 6.9. We rely on the following general construction.

Definition 6.13. Let (V, ω) be a symplectic vector space. Let W be a subspace of V. In general the restriction of ω to W will not define a symplectic structure on W because $\omega^\flat : W \to W^*$ will not be injective. However, set $Z := \{w \in W : \omega(w, \cdot)_{|W} \equiv 0\}$. Then ω *reduces* to a symplectic structure on the quotient space W/Z, defined by

$$\omega([w], [w']) := \omega(w, w').$$

In our case we can set $V := L^2(\mu)$ and $W := \overline{\operatorname{Ham} \mathcal{X}_c}^\mu$. The 2-form

$$(6.6) \qquad \hat{\Omega}_\mu(X, Y) := \int_{\mathbb{R}^{2d}} \omega(X, Y) \, d\mu$$

defines a symplectic structure on $L^2(\mu)$. The restriction of $\hat{\Omega}_\mu$ defines a 2-form

$$\bar{\Omega}_\mu : \overline{\operatorname{Ham} \mathcal{X}_c}^\mu \times \overline{\operatorname{Ham} \mathcal{X}_c}^\mu \to \mathbb{R}.$$

Notice that $\hat{\Omega}_\mu$ is continuous in the sense of Definition 4.5, so $\bar{\Omega}_\mu$ can also be defined as the unique continuous extension of the 2-form

$$(6.7) \qquad \bar{\Omega}_\mu : \operatorname{Ham} \mathcal{X}_c \times \operatorname{Ham} \mathcal{X}_c \to \mathbb{R}, \quad \bar{\Omega}_\mu(X_f, X_g) := \int \omega(X_f, X_g) \, d\mu.$$

Notice also that, for any $X \in L^2(\mu)$,

$$(6.8) \qquad \int \omega(X, X_f) \, d\mu = -\int df(X) \, d\mu = \langle div_\mu(X), f \rangle$$

so $\int \omega(X, \cdot) \, d\mu \equiv 0$ on $\overline{\operatorname{Ham} \mathcal{X}_c}^\mu$ iff $X \in \operatorname{Ker}(div_\mu)$. This calculation shows that the space Z of Definition 6.13 coincides with the space $\operatorname{Ker}(div_\mu) \cap \overline{\operatorname{Ham} \mathcal{X}_c}^\mu$. We can now define Ω_μ to be the reduced symplectic structure on the space $T_\mu \mathcal{O} = W/Z$. In terms of the identification π_μ, this yields

$$(6.9) \qquad \Omega_\mu : T_\mu \mathcal{O} \times T_\mu \mathcal{O} \to \mathbb{R}, \quad \Omega_\mu(\pi_\mu(X_f), \pi_\mu(X_g)) := \int \omega(X_f, X_g) \, d\mu.$$

Notice that Ω_μ is not necessarily continuous in the sense of Definition 4.5.

Using Equation 6.5 we can also write this as

$$\Omega_\mu(\pi_\mu(X_f), \pi_\mu(X_g)) = \int \omega(J\nabla f, J\nabla g) \, d\mu = \int g(J\nabla f, \nabla g) \, d\mu.$$

We now want to understand the geometric and differential properties of $\bar{\Omega}$. It is simple to check that $\bar{\Omega}$ is $\mathcal{G}$-invariant, in the sense that $\phi^*\bar{\Omega} = \bar{\Omega}$, for all $\phi \in \mathcal{G}$. Indeed, using Definition 4.8 and Lemma 6.3,

$$(\phi^*\bar{\Omega})_\mu(X_f, X_g) = \bar{\Omega}_{\phi_\#\mu}(\phi_*(X_f), \phi_*(X_g)) = \int_{\mathbb{R}^{2d}} \omega(X_{f\circ\phi^{-1}}, X_{g\circ\phi^{-1}}) \, d\phi_\#\mu$$

$$= \int_{\mathbb{R}^{2d}} \{f \circ \phi^{-1}, g \circ \phi^{-1}\} \, d\phi_\#\mu = \int_{\mathbb{R}^{2d}} \{f, g\} \circ \phi^{-1} \, d\phi_\#\mu$$

$$= \bar{\Omega}_\mu(X_f, X_g).$$

It then follows that Ω is also $\mathcal{G}$-invariant.

Lemma 6.14. *Given any* $X, Y, Z \in Ham\,\mathcal{X}_c$,

$$X\bar{\Omega}(Y, Z) - Y\bar{\Omega}(X, Z) + Z\bar{\Omega}(X, Y) - \bar{\Omega}([X, Y], Z) + \bar{\Omega}([X, Z], Y) - \bar{\Omega}([Y, Z], X) = 0.$$

Proof: Notice that $\bar{\Omega}(Y, Z)$ is a linear function on $\mathcal{M}$ in the sense of Example 4.6. It is thus differentiable, cf. Example 4.11, and $X\bar{\Omega}(Y, Z) = \int X\omega(Y, Z) \, d\mu$. It follows that the left hand side of the above equation reduces to $\int d\omega(X, Y, Z) \, d\mu$, which vanishes because ω is closed.

$$\text{QED.}$$

This shows that $\bar{\Omega}$ is differentiable and closed in the sense analogous to Definition 4.12, *i.e.* using Equation A.19 (or Equation A.11) with $k = 2$ instead of $k = 1$. Using the terminology of Section 4.2 we can say that $\bar{\Omega}$ is a closed pseudo linear 2-form defined on the pseudo distribution $\mu \to \overline{Ham\,\mathcal{X}_c}^\mu$ of Definition 6.9.

Remark 6.15. As in Remark 6.10, it may again be useful to emphasize a possible misconception related to the identification $\pi_\mu : \overline{Ham\,\mathcal{X}_c}^\mu/\mathrm{Ker}(div_\mu) \simeq \pi_\mu(\overline{Ham\,\mathcal{X}_c}^\mu)$. One could also restrict $\hat{\Omega}_\mu$ to the subspace $W' := \pi_\mu(\overline{Ham\,\mathcal{X}_c}^\mu)$, obtaining a 2-form

$$\Omega'_\mu(\pi_\mu(X_f), \pi_\mu(X_g)) = \int \omega(\pi_\mu(J\nabla f), \pi_\mu(J\nabla g)) \, d\mu.$$

It is important to realize that Ω'_μ does *not* coincide, under π_μ, with Ω_μ. Specifically, Ω'_μ differs from Ω_μ in that it does not take into account the divergence components of X_f, X_g.

In the framework of [5] it is more natural to work in terms of the space $\pi_\mu(\overline{Ham\,\mathcal{X}_c}^\mu) \subseteq T_\mu\mathcal{M}$ than in terms of $\overline{Ham\,\mathcal{X}_c}^\mu/\mathrm{Ker}(div_\mu)$. From this point of view, the choice of Ω_μ as a symplectic structure on $T_\mu\mathcal{O}$ may seem less natural than the choice of Ω'_μ. The fact that Ω_μ is even well-defined on $T_\mu\mathcal{O}$ follows only from Equation 6.8. Our reasons for preferring Ω_μ are based on its geometric and differential properties seen above. Together with Remark 6.10, this shows that from a symplectic viewpoint the identification π_μ is not natural.

We can now define the concept of a Hamiltonian flow on $\mathcal{M}$ as follows.

Definition 6.16. Let $F : \mathcal{M} \to \mathbb{R}$ be a differentiable function on $\mathcal{M}$ with gradient ∇F. We define the *Hamiltonian vector field* associated to F to be $X_F(\mu) := \pi_\mu(-J\nabla F)$. A *Hamiltonian flow* on $\mathcal{M}$ is a solution to the equation

$$\frac{\partial \mu_t}{\partial t} = -div_{\mu_t}(X_F).$$

Let X_F be a Hamiltonian vector field. Choose any $Y \in T_\mu \mathcal{O}$. We can then write $Y = \pi_\mu(-J\tilde{Y})$, for some $\tilde{Y} \in \overline{\nabla C_c^\infty}^\mu$. Notice that

$$\Omega_\mu(X_F, Y) = \hat{\Omega}_\mu(\nabla F, \tilde{Y}) = \hat{G}_\mu(\nabla F, -J\tilde{Y}) = \hat{G}_\mu(\nabla F, Y) = dF(Y).$$

This shows that X_F satisfies the analogue of Equation 6.2 along $\mathcal{O}$, justifying the terminology of Definition 6.16. We refer to [3] and to [25] for specific results concerning Hamiltonian flows on $\mathcal{M}$.

6.3. **Algebraic properties of the symplectic distribution.** Regardless of Remarks 6.10 and 6.15, from the point of view of [5] it is interesting to understand the linear-algebraic properties of the symplectic spaces $(T_\mu\mathcal{O}, \Omega_\mu)$, viewed as subspaces $\pi_\mu(\overline{Ham\,\mathcal{X}_c}^\mu) \subseteq T_\mu\mathcal{M}$. Throughout this section we will use this identification. We will mainly work in terms of the complex structure J on $\mathbb{R}^{2d}$ and of certain related maps. This will also serve to emphasize the role played by J within this theory. The key to this construction is of course the peculiar compatibility between the standard structures $g := \langle \cdot, \cdot \rangle$, ω and J on $\mathbb{R}^{2d}$, which we emphasize as follows.

Definition 6.17. Let V be a vector space endowed with both a metric g and a symplectic structure ω. Notice that there exists a unique injective $A \in End(V)$ such that $\omega(\cdot, \cdot) = g(A\cdot, \cdot)$. Using the isomorphism $g^\sharp : V^* \to V$ induced by g, $A = g^\sharp \circ \omega^\flat$ so A is surjective iff $\omega^\flat$ is an isomorphism.

The fact that ω is anti-symmetric implies that A is anti-selfadjoint, *i.e.* $A^* = -A$. We say that (ω, g) are *compatible* if A is an isometry, *i.e.* $A^* = A^{-1}$. In this case $A^2 = -Id$, *i.e.* A is a *complex structure* on V. A subspace $W \subseteq V$ is *symplectic* if the restriction of ω to W defines a symplectic structure on W. In particular, if g and ω are compatible than any complex subspace of V is symplectic.

The analogous definitions hold for a smooth manifold M endowed with a Riemannian metric g and a symplectic structure ω. In general, given any function f on M, the Hamiltonian vector field X_f is related to the gradient field ∇f as follows: $X_f = A^{-1}\nabla f$. If g and ω are compatible then $X_f = -A\nabla f$.

The standard structures g and ω on $\mathbb{R}^{2d}$ are of course the primary example of compatible structures. Given any $\mu \in \mathcal{M}$, $\hat{G}_\mu$ and $\hat{\Omega}_\mu$ (defined in Equations 2.4 and 6.6) are compatible structures on $L^2(\mu)$. In this case the corresponding automorphism is the isometry

$$J : L^2(\mu) \to L^2(\mu), \quad (JX)(x) := J(X(x)).$$

Example 6.18. Notice that $\overline{Ham\,\mathcal{X}_c}^\mu = -J(T_\mu\mathcal{M})$. Thus $\overline{Ham\,\mathcal{X}_c}^\mu$ is J-invariant iff $T_\mu\mathcal{M}$ is J-invariant iff $T_\mu\mathcal{O} = T_\mu\mathcal{M}$. In this case, $\bar{\Omega}_\mu = \Omega_\mu = \Omega'_\mu$. Example 6.12 shows that this is the case if μ is a Dirac measure. Example 6.12 also shows that if $\mu = \rho\mathcal{L}$ for some $\rho > 0$ then the space ∇C_c^∞ is *totally real*, *i.e.* $J(\nabla C_c^\infty) \cap \nabla C_c^\infty = Ham\,\mathcal{X}_c \cap \nabla C_c^\infty = \{0\}$.

Our first goal is to characterize the orthogonal complement of the closure of $T_\mu\mathcal{O}$ in $T_\mu\mathcal{M}$. To this end, recall that any continuous map $P : H \to H$ on a Hilbert space H satisfies $Im(P)^\perp = Ker(P^*)$, where $P^* : H \to H$ is the adjoint map. This yields an orthogonal decomposition $H = \overline{Im(P)} \oplus Ker(P^*)$.

Example 6.19. One example of such a decomposition is Decomposition 2.5, corresponding to the map $P := \pi_\mu$ defined on $H := L^2(\mu)$: in this case $Im(P)$ is closed and π_μ is self-adjoint so $Ker(P^*) = Ker(\pi_\mu)$.

Another example is provided by the map $P := \pi_\mu \circ J$, again defined on $L^2(\mu)$. In this case $P^* = -J \circ \pi_\mu$ and $\mathrm{Im}(P) = \mathrm{Im}(\pi_\mu)$, $\mathrm{Ker}(P^*) = \mathrm{Ker}(\pi_\mu)$ so the decomposition corresponding to P again coincides with Decomposition 2.5. On the other hand, $\mathrm{Im}(P^*) = -J(\mathrm{Im}(\pi_\mu))$ and $\mathrm{Ker}(P) = J^{-1}(\mathrm{Ker}(\pi_\mu)) = -J(\mathrm{Ker}(\pi_\mu))$ so the decomposition corresponding to P^* is the $(-J)$-rotation of Decomposition 2.5, *i.e.*

$$L^2(\mu) = \mathrm{Im}(P^*) \oplus \mathrm{Ker}(P) = -J(\overline{\nabla C_c^\infty}^\mu) \oplus -J(\mathrm{Ker}(div_\mu)).$$

In other words, there exists an orthogonal decomposition

(6.10) $$L^2(\mu) = \overline{Ham\, \mathcal{X}_c}^\mu \oplus Ker(\pi_\mu \circ J).$$

Lemma 6.20. *For any $\mu \in \mathcal{M}$ there exists an orthogonal decomposition*

(6.11) $$T_\mu \mathcal{M} = \overline{T_\mu \mathcal{O}} \oplus Ker((\pi_\mu \circ (-J))_{|T_\mu \mathcal{M}}).$$

In particular, the restriction of $\pi_\mu \circ (-J)$ to $\overline{T_\mu \mathcal{O}}$ is injective.

Proof: We introduce the following notation: given any map P defined on $L^2(\mu)$, let P' denote its restriction to the closed subspace $T_\mu \mathcal{M} = \mathrm{Im}(\pi_\mu)$.

Set $P := \pi_\mu \circ (-J)$. Then $\mathrm{Im}(P') \subseteq \mathrm{Im}(\pi_\mu)$ so we can think of P' as a map $P' : T_\mu \mathcal{M} \to T_\mu \mathcal{M}$, yielding a decomposition $T_\mu \mathcal{M} = \overline{\mathrm{Im}(P')} \oplus \mathrm{Ker}(P'^*)$. It is simple to check that

$$P'^* = (\pi_\mu \circ P^*)' = (\pi_\mu \circ J \circ \pi_\mu)' = (\pi_\mu \circ J)'.$$

In particular, P' is anti-selfadjoint. This implies that $\mathrm{Ker}(P'^*) = \mathrm{Ker}(P')$ so

$$T_\mu \mathcal{M} = \overline{\mathrm{Im}(P')} \oplus \mathrm{Ker}(P') = \overline{T_\mu \mathcal{O}} \oplus \mathrm{Ker}(P').$$

<div align="right">QED.</div>

It follows from Lemma 6.20 that any element $X \in T_\mu \mathcal{O}$ can be uniquely written as $X = \pi_\mu(-J\tilde{X})$, for some $\tilde{X} \in (\mathrm{Ker}((\pi_\mu \circ (-J))_{|T_\mu \mathcal{M}}))^\perp$.

Remark 6.21. Let F be a differentiable function on $\mathcal{M}$ and let X_F be the corresponding Hamiltonian vector field. It follows from Lemma 6.20 that X_F depends only on the component of ∇F tangent to $\overline{T_\mu \mathcal{O}}$. Intuitively, this suggests that the corresponding Hamiltonian flow should depend only on the restriction of F to $\mathcal{O}$.

Example 6.22. Let P' denote the restriction of $\pi_\mu \circ (-J)$ to $T_\mu \mathcal{M}$. It follows from Example 6.12 that if $\mu = \delta_x$ then the map P' is an isomorphism of $T_\mu \mathcal{M}$. If instead $\mu = \rho \mathcal{L}$ for some $\rho > 0$ then P' is neither injective nor surjective.

Lemma 6.23. *For any $\mu \in \mathcal{M}$ the map*

$$\Omega_\mu^\flat : T_\mu \mathcal{O} \to (T_\mu \mathcal{O})^*, \quad X \mapsto \Omega_\mu(X, \cdot)$$

is a (non-continuous) isomorphism.

Proof: Given any $X, Y \in T_\mu \mathcal{O}$, we can write $X = \pi_\mu(-J\tilde{X})$, $Y = \pi_\mu(-J\tilde{Y})$ for some $\tilde{X}, \tilde{Y} \in T_\mu \mathcal{M}$. Then

$$\Omega_\mu(X, Y) = \hat{\Omega}_\mu(\tilde{X}, \tilde{Y}) = \hat{G}_\mu(\tilde{X}, -J\tilde{Y}) = \hat{G}_\mu(\tilde{X}, Y) \leq \|\tilde{X}\|_\mu \cdot \|Y\|_\mu.$$

This proves that $\Omega_\mu^\flat$ is well-defined, *i.e.* that $\Omega_\mu^\flat(X) \in (T_\mu \mathcal{O})^*$.

Assume $\Omega_\mu(X, Y) = 0$ for all $Y \in T_\mu \mathcal{O}$. Then, as above, $\hat{G}_\mu(\tilde{X}, Y) = 0$ for all $Y \in T_\mu \mathcal{O}$ so $\tilde{X} \in (T_\mu \mathcal{O})^\perp$. It follows from Lemma 6.20 that $X = 0$ so $\Omega_\mu^\flat$ is injective.

Now choose $\Lambda \in (T_\mu\mathcal{O})^*$. Since $(T_\mu\mathcal{O})^* = (\overline{T_\mu\mathcal{O}})^*$, there exists $\tilde{X} \in \overline{T_\mu\mathcal{O}}$ such that $\Lambda = \hat{G}_\mu(\tilde{X}, \cdot)$. Set $X := \pi_\mu(-J\tilde{X}) \in T_\mu\mathcal{O}$. Then

$$\Lambda(Y) = \hat{G}_\mu(\tilde{X}, \pi_\mu(-J\tilde{Y})) = \hat{G}_\mu(\tilde{X}, -J\tilde{Y}) = \hat{\Omega}_\mu(\tilde{X}, \tilde{Y}) = \Omega_\mu(X, Y).$$

This shows that $\Omega_\mu^\flat$ is surjective. QED.

Remark 6.24. Lemma 6.20 shows that the map $\pi_\mu \circ (-J) : \overline{T_\mu\mathcal{O}} \to T_\mu\mathcal{O}$ is invertible. Its inverse is related to the notions introduced in Definition 6.17 as follows. We can use the isomorphism $(\overline{T_\mu\mathcal{O}})^* \simeq \overline{T_\mu\mathcal{O}}$ induced by $\hat{G}_\mu$ to define a (non-continuous) isomorphism $A : T_\mu\mathcal{O} \to \overline{T_\mu\mathcal{O}}$ such that $\Omega_\mu(\cdot, \cdot) = \hat{G}_\mu(A\cdot, \cdot)$ on $T_\mu\mathcal{O}$. Notice that

$$\hat{G}_\mu(AX, Y) = \Omega_\mu(X, Y) = \hat{\Omega}_\mu(\tilde{X}, \tilde{Y}) = \hat{G}_\mu(\tilde{X}, Y),$$

where we use the notation introduced in the proof of Lemma 6.23. This shows that $AX = \tilde{X}$ so $A^{-1} = \pi_\mu \circ (-J)$.

If μ is a Dirac measure it is clearly the case that G_μ and Ω_μ are a compatible pair in the sense of Definition 6.17. This amounts to stating that $(\pi_\mu \circ (-J))^2 = -Id$ on $T_\mu\mathcal{O}$. It is not clear if this is true in general, even under the additional assumption that $T_\mu\mathcal{O} = \overline{T_\mu\mathcal{O}}$.

7. THE SYMPLECTIC FOLIATION AS A POISSON STRUCTURE

Most naturally occurring symplectic foliations owe their existence to an underlying *Poisson structure*. The symplectic foliation described in Section 6.2 is no exception. The existence of a Poisson structure on a certain space of distributions was pointed out in [33]. It stems from the fact that the symplectic structure on $\mathbb{R}^{2d}$ adds new structure into the framework of Section 3.3. The goal of this section is to show that, reduced to $\mathcal{M}$, this Poisson structure coincides with the symplectic structure Ω defined in Section 6.2. We achieve this in Section 7.2 after briefly reviewing the relevant notions. We refer to [32] for many details.

7.1. **Review of Poisson geometry.**
Recall from Section 6.1 that any symplectic structure ω on a manifold M induces a Lie bracket on the space of functions $C^\infty(M)$. Using the Liebniz rule for the derivative of the product of two functions, we see that the corresponding operators $\{\cdot, h\}$ have the following property:

$$\{fg, h\} = d(fg)(X_h) = df(X_h)g + dg(X_h)f = \{f, h\}g + \{g, h\}f.$$

This leads to the following natural "weakening" of Symplectic Geometry.

Definition 7.1. Let M be a smooth manifold. A *Poisson structure* on M is a Lie bracket $\{\cdot, \cdot\}$ on $C^\infty(M)$ such that each operator $\{\cdot, h\}$ is a *derivation* on functions, *i.e.*

$$\{fg, h\} = \{f, h\}g + \{g, h\}f.$$

A *Poisson manifold* is a manifold endowed with a Poisson structure.

On any finite-dimensional manifold it is known that the space of derivations on functions is isomorphic to the space of vector fields. Thus on any Poisson manifold any function h defines a vector field which we denote X_h: it is uniquely defined by the property that

$$df(X_h) = \{f, h\}, \quad \forall f \in C^\infty(M).$$

We call X_h the *Hamiltonian vector field* defined by h. As in Section 6.1, this process defines a map

(7.1) $$C^\infty(M) \to \mathcal{X}(M), \quad f \mapsto X_f.$$

The kernel of this map includes the space of constant functions, but in general it will be larger. We call these the *Casimir functions* of the Poisson manifold. Its image defines the space $\mathrm{Ham}(M)$ of *Hamiltonian vector fields*. Lemma 6.4 applies with the same proof to show that the map of Equation 7.1 is a Lie algebra homomorphism (up to sign).

At each point $x \in M$, the set of Hamiltonian vector fields evaluated at that point define a subspace of $T_x M$. The union of such subspaces is known as the *characteristic distribution* of the Poisson manifold. This distribution is *integrable* in the sense that M admits a smooth foliation such that each subspace is the tangent space of the corresponding leaf. In particular each leaf has a well-defined dimension, but this dimension may vary from leaf to leaf. Each leaf has a symplectic structure defined by setting

(7.2) $$\omega(X_f, X_g) := \{f, g\}.$$

Remark 7.2. Notice that for any given Hamiltonian vector field X_f, the corresponding function f is well-defined only up to Casimir functions. It is however simple to check that ω is a well-defined 2-form on each leaf, *i.e.* it is independent of the particular choices made for f and g. It is also non-degenerate. The fact that ω is closed follows from the Jacobi identity for $\{\cdot, \cdot\}$.

Remark 7.3. Notice that the definition of a Poisson manifold does not include a metric. Thus there is in general no intrinsic way to extend ω to a 2-form on M.

The following result is standard.

Proposition 7.4. *Any Poisson manifold admits a symplectic foliation, of varying rank. Each leaf is preserved by the flow of any Hamiltonian vector field. Any Casimir function restricts to a constant along any leaf of the foliation.*

Poisson manifolds are of interest in Mechanics because they provide the following generalization of the standard symplectic notion of Hamiltonian flows.

Definition 7.5. A *Hamiltonian flow* on M is a solution of the equation $d/dt(x_t) = X_f(x_t)$, for some function f on M.

It follows from Proposition 7.4 that if the initial data belongs to a specific leaf, then the corresponding Hamiltonian flow is completely contained within that leaf. It is simple to check that if x_t is Hamiltonian then f is constant along x_t.

The theory of Lie algebras provides one of the primary classes of examples of Poisson manifolds. To explain this we introduce the following notation, once again restricting our attention to the finite-dimensional case. Let V be a finite-dimensional vector space, whose generic element will be denoted v. Let V^* be its dual, with generic element ϕ. Let V^{**} be the bidual space, defined as the space of linear maps $V^* \to \mathbb{R}$. We will think of this as a subspace of the space of smooth maps on V^*, with generic element $f = f(\phi)$. We can identify V with V^{**} via the map

(7.3) $$V \to V^{**}, \quad v \mapsto f_v \text{ where } f_v(\phi) := \phi(v).$$

Now assume V is a Lie algebra. We will write $V = \mathfrak{g}$. Consider the vector space $\mathfrak{g}^*$ dual to $\mathfrak{g}$. We want to show that the Lie algebra structure on $\mathfrak{g}$ induces a natural Poisson structure

on $\mathfrak{g}^*$. Let f be a smooth function on $\mathfrak{g}^*$. Its linearization at ϕ is an element of the bidual: $df_{|\phi} \in \mathfrak{g}^{**}$. It thus corresponds via the map of Equation 7.3 to an element $\delta f/\delta \phi_{|\phi} \in \mathfrak{g}$. We can now define a Lie bracket on $\mathfrak{g}^*$ by setting:

$$(7.4) \qquad \{f,g\}(\phi) := \phi([\delta f/\delta \phi_{|\phi}, \delta g/\delta \phi_{|\phi}]),$$

where $[\cdot, \cdot]$ denotes the Lie bracket on $\mathfrak{g}$. One can show that this operation satisfies the Jacobi identity and defines a Poisson structure on $\mathfrak{g}^*$.

Example 7.6. Assume f is a linear function on $\mathfrak{g}^*$, $f = f_v$ (as in Equation 7.3). Then $\delta f/\delta \phi \equiv v$, so $\{f_v, f_w\}(\phi) = \phi([v, w])$.

We now want to characterize the Hamiltonian vector fields and symplectic leaves of $\mathfrak{g}^*$. Unsurprisingly, this is best done in terms of Lie algebra theory. Every finite-dimensional Lie algebra is the Lie algebra of a (unique connected and simply connected) Lie group G. Recall from Section A.2 the adjoint representation of G on $\mathfrak{g}$,

$$G \to Aut(\mathfrak{g}), \quad g \mapsto Ad_g.$$

Differentiating this defines the *adjoint representation* of $\mathfrak{g}$ on $\mathfrak{g}$,

$$(7.5) \qquad ad : \mathfrak{g} \to End(\mathfrak{g}), \quad v = d/dt(g_t)_{|t=0} \mapsto ad_v := d/dt(Ad_{g_t})_{|t=0}.$$

It follows from Lemma A.12 that $ad_v(w) = [v, w]$.

By duality we obtain the *coadjoint representation* of G on $\mathfrak{g}^*$,

$$G \to Aut(\mathfrak{g}^*), \quad g \mapsto (Ad_{g^{-1}})^*.$$

Notice that once again we have used inversion to ensure that this remains a group homomorphism, cf. Remark A.9. We can differentiate this to obtain the *coadjoint representation* of $\mathfrak{g}$ on $\mathfrak{g}^*$, which can also be written in terms of the duals of the maps in Equation 7.5:

$$(7.6) \qquad ad^* : \mathfrak{g} \to End(\mathfrak{g}^*), \quad v \mapsto (-ad_v)^*.$$

The following result is standard, cf. *e.g.* [32] Formula 10.7.4.

Lemma 7.7. *The Hamiltonian vector field corresponding to a smooth function f on $\mathfrak{g}^*$ is*

$$X_f(\phi) := (-ad_{\delta f/\delta \phi_{|\phi}})^*(\phi).$$

Thus the leaves of the symplectic foliation of $\mathfrak{g}^*$ are the orbits of the coadjoint representation.

7.2. **The symplectic foliation of $\mathcal{M}$, revisited.** Following [33] we now apply the ideas of Section 7.1 to the case where $\mathfrak{g}$ is the Lie algebra of $\mathrm{Ham}_c(\mathbb{R}^{2d})$. Since this is an infinite-dimensional algebra, the following discussion will be purely formal.

We saw in Remark 6.2 that $\mathfrak{g}$ can be identified with the space of compactly-supported functions:

$$(7.7) \qquad C_c^\infty(\mathbb{R}^{2d}) \simeq \mathrm{Ham}\,\mathcal{X}_c(\mathbb{R}^{2d}), \quad f \mapsto X_f.$$

Its dual is then the distribution space $(C_c^\infty)^*$. Section 7.1 suggests that $(C_c^\infty)^*$ has a canonical Poisson structure, defined as in Equation 7.4. We can identify the Poisson bracket, Hamiltonian vector fields and symplectic leaves on $(C_c^\infty)^*$ very explicitly, as follows.

For simplicity let us restrict our attention to the linear functions on $(C_c^\infty)^*$ defined by functions $f \in C_c^\infty$ as follows:

$$(7.8) \qquad F_f : (C_c^\infty)^* \to \mathbb{R}, \quad F_f(\mu) := \langle \mu, f \rangle.$$

Example 7.6 shows that the Poisson bracket of two such functions F_f and F_g can be written in terms of the Lie bracket on C_c^∞:

$$(7.9) \qquad \{F_f, F_g\}_{(C_c^\infty)^*}(\mu) = \langle \mu, \{f, g\}_{\mathbb{R}^{2d}} \rangle = \langle \mu, \omega(X_f, X_g) \rangle.$$

Lemma 7.7 gives an explicit formula for the corresponding Hamiltonian vector fields X_{F_f}: at $\mu \in (C_c^\infty)^*$, $X_{F_f}(\mu) \in T_\mu(C_c^\infty)^* = (C_c^\infty)^*$ is given by

$$\langle X_{F_f}(\mu), g \rangle = \langle (-ad_f)^*(\mu), g \rangle = \langle \mu, -ad_f(g) \rangle = \langle \mu, -\{f, g\}_{\mathbb{R}^{2d}} \rangle = \langle \mu, dg(X_f) \rangle$$
$$= -\langle div_\mu(X_f), g \rangle.$$

In other words, $X_{F_f}(\mu) = -div_\mu(X_f)$.

Lemma 7.7 also shows that the leaves of the symplectic foliation are the orbits of the coadjoint representation of $\mathrm{Ham}_c(\mathbb{R}^{2d})$ on $(C_c^\infty)^*$. Let us identify the coadjoint representation explicitly. Recall from Lemma A.19 that the adjoint representation of $\mathrm{Ham}_c(\mathbb{R}^{2d})$ on $\mathrm{Ham}\,\mathcal{X}_c$ is the push-forward operation. Lemma 6.3 shows that, under the isomorphism of Equation 7.7, push-forward becomes composition. Thus the adjoint representation of $\mathrm{Ham}_c(\mathbb{R}^{2d})$ on $\mathrm{Ham}\,\mathcal{X}_c$ corresponds to the following representation of $\mathrm{Ham}_c(\mathbb{R}^{2d})$ on $C_c^\infty(\mathbb{R}^{2d})$:

$$(7.10) \qquad Ad : \mathrm{Ham}_c(\mathbb{R}^{2d}) \to Aut(C_c^\infty(\mathbb{R}^{2d})), \quad Ad_\phi(f) := f \circ \phi^{-1}.$$

The following calculation then shows that the coadjoint representation of $\mathrm{Ham}_c(\mathbb{R}^{2d})$ on $(C_c^\infty)^*$ is simply the natural action of $\mathrm{Ham}_c(\mathbb{R}^{2d})$ introduced in Section 3.3:

$$\langle (Ad_{\phi^{-1}})^*(\mu), f \rangle = \langle \mu, Ad_{\phi^{-1}}(f) \rangle = \langle \mu, f \circ \phi \rangle = \langle \phi \cdot \mu, f \rangle.$$

The symplectic structure on each leaf is given by Equation 7.2:

$$(7.11) \qquad \omega_\mu(-div_\mu(X_f), -div_\mu(X_g)) := \{F_f, F_g\}_{(C_c^\infty)^*}(\mu) = \langle \mu, \omega(X_f, X_g) \rangle.$$

Remark 7.8. Notice that Poisson brackets and Hamiltonian vector fields are of first order with respect to the functions involved. We can use this fact to reduce the study of general functions $F : (C_c^\infty)^* \to \mathbb{R}$ to the study of linear functions on $(C_c^\infty)^*$, as presented above. For example if $\nabla_\mu F = \nabla_\mu F_f$, for some linear F_f as above, then $X_F(\mu) = X_{F_f}(\mu)$.

Let us now restrict our attention to $\mathcal{M} \subset (C_c^\infty)^*$. We want to show that the data defined by the Poisson structure on $(C_c^\infty)^*$ restricts to the objects defined in Section 6.2. Firstly, $\mathcal{M}$ is $\mathrm{Ham}_c(\mathbb{R}^{2d})$-invariant and the action of $\mathrm{Ham}_c(\mathbb{R}^{2d})$ on $(C_c^\infty)^*$ restricts to the standard push-forward action on $\mathcal{M}$. This shows that the leaves defined above, passing through $\mathcal{M}$, coincide with the $\mathcal{G}$-orbits of Section 6.2. Now recall from Section 3.3 that, given $\mu \in \mathcal{M}$, the operator $-div_\mu$ is the natural isomorphism relating the tangent planes of Definition 2.5 to the tangent planes of $\mathcal{M} \subset (C_c^\infty)^*$. Equation 7.11 can thus be re-written as

$$\omega_\mu(\pi_\mu(X_f), \pi_\mu(X_g)) := \int_{\mathbb{R}^{2d}} \omega(X_f, X_g)\, d\mu,$$

showing that the symplectic structure defined this way coincides with the symplectic form Ω_μ defined in Equation 6.9. To conclude, we want to show that the Hamiltonian vector fields introduced in Definition 6.16 formally coincide with the Hamiltonian vector fields of the restricted Poisson structure. Let $F : \mathcal{M} \to \mathbb{R}$ be a differentiable function on $\mathcal{M}$. Fix $\mu \in \mathcal{M}$. Up to L^2_μ-closure, we can assume that $\nabla_\mu F = \nabla f$, for some $f \in C_c^\infty(\mathbb{R}^{2d})$. Example 4.11 shows that $\nabla f = \nabla_\mu F_f$, where F_f is the linear function defined in Equation 7.8. Using

Remark 7.8, the Hamiltonian vector of F at μ defined by the Poisson structure is thus $X_F(\mu) = X_{F_f}(\mu) = -div_\mu(X_f)$. In terms of the tangent space $T_\mu\mathcal{M}$, we can write this as

$$(7.12) \qquad X_F(\mu) = \pi_\mu(X_f) = \pi_\mu(-J\nabla f) = \pi_\mu(-J\nabla_\mu F).$$

It thus coincides with the vector field given in Definition 6.16.

Remark 7.9. The identification of $(C_c^\infty)^*$ with the dual Lie algebra of $\mathrm{Ham}_c(\mathbb{R}^{2d})$ relied on the normalization introduced in Remark 6.2. In turn, this was based on our choice to restrict our attention to diffeomorphisms with compact support. In some situations one might want to relax this assumption. This would generally mean losing the possibility of a normalization so Equation 6.3 would leave us only with an identification $\mathrm{Ham}\,\mathcal{X} \simeq C^\infty(M)/\mathbb{R}$. Dualizing this space would then, roughly speaking, yield the space of measures of integral zero: we would thus get a Poisson structure on this space but not on $\mathcal{M}$. However this issue is purely technical and can be avoided by changing Lie group, as follows.

Consider the group G of diffeomorphisms on $\mathbb{R}^{2d} \times \mathbb{R}$ preserving the *contact form* $dz - y^i dx^i$. It can be shown that its Lie algebra is isomorphic to the space of functions on $\mathbb{R}^{2d} \times \mathbb{R}$ which are constant with respect to the new variable z: it is thus isomorphic to the space of functions on $\mathbb{R}^{2d}$, so the dual Lie algebra is, roughly, the space of measures on $\mathbb{R}^{2d}$; in particular, it contains $\mathcal{M}$ as a subset. This group has a one-dimensional center $Z \simeq \mathbb{R}$, defined by translations with respect to z. The center acts trivially in the adjoint and coadjoint representations, so the coadjoint representation reduces to a representation of the group G/Z, which one can show to be isomorphic to the group of Hamiltonian diffeomorphisms of $\mathbb{R}^{2d}$. The coadjoint representation of G reduces to the standard push-forward action of Hamiltonian diffeomorphisms, and the theory can now proceed as before.

Appendix A. Review of relevant notions of Differential Geometry

The goal of the first two sections of this appendix is to summarize standard facts concerning Lie groups and calculus on finite-dimensional manifolds, thus laying out the terminology, notation and conventions which we use throughout this paper. We refer to [26] and [32] for details. The third section introduces the notion of invariant cohomology. The point of view adopted here might be new. It provides useful analogies for the notion of "pseudo forms" introduced in Section 4.2. The fourth section provides some basic facts concerning the infinite-dimensional Lie groups relevant to this paper.

A.1. **Calculus of vector fields and differential forms.** Let M be a connected differentiable manifold of dimension D, not necessarily compact. Let $\mathrm{Diff}(M)$ denote the group of diffeomorphisms of M. Let $C^\infty(M)$ denote the space of smooth functions on M. Let TM denote the tangent bundle of M and $\mathcal{X}(M)$ the corresponding space of sections, *i.e.* the space of smooth vector fields. Let T^*M denote the cotangent bundle of M. To simplify notation, $\Lambda^k M$ will denote both the bundle of k-forms on M and the space of its sections, *i.e* the space of smooth k-forms on M. Notice that $\Lambda^0(M) = C^\infty(M)$ and $\Lambda^1 M = T^*M$ (or the space of smooth 1-forms).

Let $\phi \in \mathrm{Diff}(M)$. Taking its differential yields linear maps

$$(A.1) \qquad \nabla\phi : T_x M \to T_{\phi(x)} M, \quad v \mapsto \nabla\phi \cdot v,$$

thus a bundle map which we denote $\nabla\phi : TM \to TM$. We will call $\nabla\phi$ the *lift* of ϕ to TM.

By duality we obtain linear maps

$$(\nabla\phi)^* : T^*_{\phi(x)}M \to T^*_x M, \quad \alpha \mapsto \alpha \circ \nabla\phi,$$

and more generally k-multilinear maps

(A.2) $$(\nabla\phi)^* : \Lambda^k_{\phi(x)}M \to \Lambda^k_x M, \quad \alpha \mapsto \alpha(\nabla\phi\,\cdot,\dots,\nabla\phi\,\cdot).$$

This defines bundle maps $(\nabla\phi)^* : \Lambda^k M \to \Lambda^k M$ which we call the *lift* of ϕ to $\Lambda^k M$.

Remark A.1. Notice the different behaviour under composition of diffeomorphisms: $\nabla(\phi \circ \psi) = \nabla\phi \circ \nabla\psi$ while $(\nabla(\phi \circ \psi))^* = (\nabla\psi)^* \circ (\nabla\phi)^*$. We will take this into account and generalize it in Section A.2 via the notion of left versus right group actions.

We can of course apply these lifted maps to sections of the corresponding bundles. In doing so one needs to ensure that the correct relationship between $T_x M$ and $T_{\phi(x)}M$ is maintained; we emphasize this with a change of notation, as follows.

The *push-forward* operation on vector fields is defined by

(A.3) $$\phi_* : \mathcal{X}(M) \to \mathcal{X}(M), \quad \phi_* X := (\nabla\phi \cdot X) \circ \phi^{-1}.$$

The corresponding operation on k-forms is the *pull-back*, defined by

(A.4) $$\phi^* : \Lambda^k(M) \to \Lambda^k(M), \quad \phi^*\alpha := ((\nabla\phi)^*\alpha) \circ \phi.$$

Definition A.2. Let V be a vector space. A bilinear anti-symmetric operation

$$V \times V \to V, \quad (v, w) \mapsto [v, w]$$

is a *Lie bracket* if it satisfies the *Jacobi identity*

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0.$$

A *Lie algebra* is a vector space endowed with a Lie bracket.

The space of smooth vector fields has a natural Lie bracket. Given two vector fields X, Y on M, we define $[X, Y]$ in local coordinates as follows:

$$[X, Y] := \nabla Y \cdot X - \nabla X \cdot Y.$$

It is simple to show that this operation indeed satisfies the Jacobi identity. Let ϕ_t denote the *flow* of X on $\mathbb{R}^D$, *i.e.* the 1-parameter group of diffeomorphisms obtained by integrating X as follows:

(A.5) $$d/dt(\phi_t(x)) = X(\phi_t(x)), \quad \phi_0(x) = x.$$

It is then simple to check that

(A.6) $$[X, Y] = -d/dt(\phi_{t*}Y)_{|t=0} = d/dt(\phi_{-t*}Y)_{|t=0} = d/dt((\phi_t^{-1})_* Y)_{|t=0}.$$

Equation A.6 gives a coordinate-free expression for the Lie bracket. It also suggests an analogous operation for more general tensor fields. We will restrict our attention to the case of differential forms.

Let α be a smooth k-form on M. Let X, ϕ_t be as above. We define the *Lie derivative* of α in the direction of X to be the k-form defined as follows:

(A.7) $$\mathcal{L}_X \alpha := d/dt(\phi_t^* \alpha)_{|t=0}.$$

The fact that $t \mapsto \phi_t$ is a homomorphism leads to the fact that $d/dt(\phi_t^*\alpha)_{|t=t_0} = \phi_{t_0}^*(\mathcal{L}_X\alpha)$. Thus $\mathcal{L}_X\alpha \equiv 0$ if and only if $\phi_t^*\alpha \equiv \alpha$, *i.e.* ϕ_t preserves α. This can be generalized to time-dependent vector fields as follows.

Lemma A.3. *Let X_t be a t-dependent vector field on M. Let $\phi_t = \phi_t(x)$ be its flow, defined by*

(A.8) $$d/dt(\phi_t(x)) = X_t(\phi_t(x)), \quad \phi_0(x) = x.$$

Let α be a k-form on M. Then $d/dt(\phi_t^*\alpha)_{|t_0} = \phi_{t_0}^*(\mathcal{L}_{X_{t_0}}\alpha)$. In particular, $\phi_t^*\alpha \equiv \alpha$ iff $\mathcal{L}_{X_t}\alpha \equiv 0$.

Proof: For any fixed s, let ψ_t^s be the flow of X_s, *i.e.*

$$d/dt(\psi_t^s(x)) = X_s(\psi_t^s(x)), \quad \psi_0^s(x) = x.$$

Then $\psi_t^{t_0} \circ \phi_{t_0}(x)$ satisfies

$$d/dt(\psi_t^{t_0} \circ \phi_{t_0}(x))_{|t=0} = X_{t_0}(\phi_{t_0}(x)), \quad \psi_0^{t_0} \circ \phi_{t_0}(x) = \phi_{t_0}(x)$$

so $\psi_t^{t_0} \circ \phi_{t_0}(x)$ at $t = 0$ and ϕ_t at $t = t_0$ coincide up to first order, showing that

$$d/dt(\phi_t^*\alpha)_{|t=t_0} = d/dt((\psi_t^{t_0} \circ \phi_{t_0})^*\alpha)_{|t=0} = \phi_{t_0}^*(d/dt((\psi_t^{t_0})^*\alpha)_{|t=0}) = \phi_{t_0}^*(\mathcal{L}_{X_{t_0}}\alpha).$$

$$\text{QED.}$$

Notice that if we define $\phi^*Y := (\phi^{-1})_*Y$ and we define $\mathcal{L}_X Y := d/dt(\phi_t^*Y)_{|t=0}$, then Equation A.6 shows that $\mathcal{L}_X Y = [X, Y]$.

Remark A.4. Various formulae relate the above operations, leading to quick proofs of useful facts. For example, the fact

(A.9) $$\mathcal{L}_{[X,Y]}\alpha = \mathcal{L}_X(\mathcal{L}_Y\alpha) - \mathcal{L}_Y(\mathcal{L}_X\alpha)$$

shows that if the flows of X and Y preserve α then so does the flow of $[X, Y]$. Also,

(A.10) $$\phi^*\mathcal{L}_X\alpha = \mathcal{L}_{\phi^*X}\phi^*\alpha.$$

Remark A.5. Notice that $\mathcal{L}_X Y$ is not a "proper" directional derivative in the sense that it is of first order also in the vector field X. In general the same is true for the Lie derivative of any tensor. The case of 0-forms, *i.e.* functions, is an exception. In this case $\mathcal{L}_X f = df(X)$ is of order zero in X and coincides with the usual notion of directional derivative. We will often simplify the notation by writing it as Xf.

We now want to introduce the *exterior differentiation* operator on smooth forms. Let α be a k-form on M. Fix any point $x \in M$ and tangent vectors $X_0, \ldots, X_k \in T_xM$. Choose any extension of each X_j to a global vector field which we will continue to denote X_j. Then, at x,

(A.11) $$d\alpha(X_0, \ldots, X_k) := \sum_{j=0}^{k}(-1)^j X_j\alpha(X_0, \ldots, \hat{X}_j, \ldots, X_k)$$
$$+ \sum_{j<l}(-1)^{j+l}\alpha([X_j, X_l], X_0, \ldots, \hat{X}_j, \ldots, \hat{X}_l, \ldots, X_k)$$

where on the right hand side the superscript ˆ denotes an omitted term and we adopt the notation for directional derivatives introduced in Remark A.5.

Lemma A.6. *$d\alpha$ is a well-defined (k+1)-form, i.e. at any point $x \in M$ it is independent of the choice of the extension. Exterior differentiation defines a first-order linear operator*

(A.12) $$d : \Lambda^k M \to \Lambda^{k+1} M$$

satisfying $d \circ d = 0$.

Remark A.7. It is not clear from the above definition that $d\alpha$ is *tensorial* in $X_0, \ldots, X_k$, *i.e.* that it is independent of the choice of extensions. The point is that cancelling occurs to eliminate the first derivatives of X_j which appear in Equation A.11. This is the main content of Lemma A.6, which is proved by showing that Equation A.11 is equivalent to the usual, local-coordinate, definition of $d\alpha$. For example, let $\alpha = \sum_{i=1}^{D} \alpha_i(x) dx^i$ be a smooth 1-form on $\mathbb{R}^D$. Then $d\alpha = \sum_{j<i} \left(\frac{\partial \alpha_i}{\partial x^j} - \frac{\partial \alpha_j}{\partial x^i} \right) dx^j \wedge dx^i$. If we identify α with the vector field $x \to (\alpha_1(x), \cdots, \alpha_D(x))^T$ then $d\alpha(X, Y) = \langle (\nabla\alpha - \nabla\alpha^T)X, Y \rangle$.

Given a k-form α and a vector field X, let $i_X \alpha$ denote the (k-1)-form $\alpha(X, \cdot, \ldots, \cdot)$ obtained by *contraction*. Then the Lie derivative and exterior differentiation are related by *Cartan's formula*:

(A.13) $$\mathcal{L}_X \alpha = d\,i_X \alpha + i_X d\alpha.$$

A.2. **Lie groups and group actions.** Recall that a group G is a *Lie group* if it has the structure of a smooth manifold and group multiplication (respectively, inversion) defines a smooth map $G \times G \to G$ (respectively, $G \to G$). We denote by e the identity element of G.

Definition A.8. We say that G has a *left action* or *acts on the left* or, more simply, *acts* on a smooth manifold M if there is a smooth map

$$G \times M \to M, \quad (g, x) \mapsto g \cdot x$$

such that $g \cdot (h \cdot x) = (gh) \cdot x$. To simplify the notation we will often write gx rather than $g \cdot x$. It is simple to see that if G has a left action on M then every $g \in G$ defines a diffeomorphism of M. More specifically, the action defines a group homomorphism $G \to \text{Diff}(M)$.

We say that G has a *right action* or *acts on the right* on M if the opposite composition rule holds: $g \cdot (h \cdot x) = hg \cdot x$. In this case it is standard to change the notation, writing $x \cdot g$ rather than $g \cdot x$: this makes the composition rule seem more natural but does not affect the substance of the definition, *i.e.* the fact that the induced map $G \to \text{Diff}(M)$ is now a group anti-homomorphism.

Remark A.9. Notice that any left action induces a natural right action as follows: $x \cdot g := g^{-1} \cdot x$. Conversely, any right action induces a natural left action: $g \cdot x := x \cdot g^{-1}$.

For any group action we can repeat the constructions of Equations A.1 and A.2. For example a left action of G on M induces a *lifted left action* of G on TM as follows:

$$G \times TM \to TM, \quad g(x, v) := (gx, \nabla g \cdot v).$$

However, we need to apply the trick introduced in Remark A.9 to obtain a coherent lifted action on T^*M or $\Lambda^k M$. For example we can define a *lifted left action* by setting

$$G \times \Lambda^k M \to \Lambda^k M, \quad g(x, \alpha) := (gx, (\nabla g^{-1})^* \alpha)$$

or a *lifted right action* by setting

$$G \times \Lambda^k M \to \Lambda^k M, \quad g(x, \alpha) := (g^{-1}x, (\nabla g)^* \alpha).$$

We can also repeat the constructions of Equations A.3 and A.4. We thus find an induced action of G on vector fields, defined by

$$(A.14) \qquad G \times \mathcal{X}(M) \to \mathcal{X}(M), \quad g \cdot X := g_* X.$$

Likewise, there is an induced action of G on k-forms. On the other hand, with respect to Section A.1 there now exists a new operation, as follows. Choose $v = d/dt(g_t)_{|t=0} \in T_e G$. For any $x \in M$ we can define the tangent vector $\tilde{v}(x) := d/dt(g_t \cdot x)_{|t=0}$. This defines a global vector field $\tilde{v}$ on M, called the *fundamental vector field* associated to v. We have thus built a map $T_e G \to \mathcal{X}(M)$.

Let us now specialize to the case $M = G$. Any Lie group G admits two natural left actions on itself. Studying these actions leads to a deeper understanding of the geometry of Lie groups, thus of group actions. The first action is given by *left translations*, as follows:

$$L : G \times G \to G, \quad (g, h) \mapsto L_g(h) := gh.$$

Let $e \in G$ denote the identity element. Fix $v = d/dt(g_t)_{|t=0} \in T_e G$. The differential ∇L_g maps $T_e G$ to $T_g G$. We may thus define a global vector field X_v on G by setting $X_v(g) := \nabla L_g \cdot v = d/dt(gg_t)_{|t=0}$. This vector field has the property of being *left-invariant* with respect to the action of G, *i.e.* $L_{g*} X_v = X_v$. Viceversa, any left-invariant vector field arises this way.

Lemma A.10. *The set of left-invariant vector fields is a finite dimensional vector space isomorphic to $T_e G$. The Lie bracket of left-invariant vector fields is a left-invariant vector field.*

It follows from Lemma A.10 that $T_e G$ admits a natural operation $[v, w]$ such that $X_{[v,w]} = [X_v, X_w]$. It follows from the Jacobi identity on vector fields that $T_e G$ equipped with this structure is a Lie algebra: we call it the *Lie algebra of G* and denote it by $\mathfrak{g}$.

Remark A.11. Given any $v \in T_e G$, we have now defined two constructions of a global vector field on G associated to v: the fundamental vector field $\tilde{v}$ and the left-invariant vector field X_v. It is simple to check that $\tilde{v}$ is invariant with respect to the *right translations*

$$R : G \times G \to G, \quad (g, h) \mapsto R_g(h) := hg.$$

This implies that the space of fundamental vector fields coincides with the space of right-invariant vector fields. The analogue of Lemma A.10 holds for right-invariant fields and can be used to define a second Lie bracket on $T_e G$. It can be checked that this new bracket is simply the negative of the old one, *i.e.* the two brackets differ only by sign.

The second action of G on itself is the *adjoint action* defined by the *inner automorphisms* $I_g(h) := ghg^{-1}$. Each of these fixes the identity and thus defines a map

$$(A.15) \qquad Ad_g := \nabla I_g : T_e G \to T_e G,$$

i.e. an automorphism of $T_e G$. In other words the adjoint action of G on G induces a left action of G on $T_e G$ called the *adjoint representation* of G.

The adjoint representation of G provides a useful way to calculate Lie brackets on $\mathfrak{g}$, as follows.

Lemma A.12. *Fix $v, w \in \mathfrak{g}$. Assume $v = d/dt(g_t)_{|t=0}$ for some $g_t \in G$. Then $[v, w] = d/dt(Ad_{g_t}(w))_{|t=0}$.*

Proof: Assume $w = d/ds(h_s)_{|s=0}$. By definition,

(A.16) $$d/dt(Ad_{g_t}(w))_{|t=0} = d/dt \, d/ds(g_t h_s g_t^{-1})_{|t,s=0}.$$

Notice that

(A.17) $$X_v(g) = \nabla L_g(v) = d/dt(gg_t)_{|t=0} = d/dt(R_{g_t}(g))_{|t=0}.$$

In particular this shows that, for $t = 0$, R_{g_t} coincides with the flow of X_v up to first order. Thus

$$[v,w] = (\mathcal{L}_{X_v} X_w)_{|e} = d/dt((R_{g_t})^* X_w)_{|e;\, t=0} = d/dt((R_{g_t^{-1}})_* X_w)_{|e;\, t=0}$$
$$= d/dt((\nabla R_{g_t^{-1}})_{|g_t} X_{w|g_t})_{|t=0} = d/dt((\nabla R_{g_t^{-1}})_{|g_t} d/ds(g_t h_s)_{|s=0})_{t=0}$$
$$= d/dt \, d/ds(g_t h_s g_t^{-1})_{|s,t=0}.$$

<div align="right">QED.</div>

Remark A.13. It is sometimes useful to distinguish the vector space $T_e G$ from the Lie algebra $\mathfrak{g}$, so as to distinguish between maps or constructions which involve the Lie bracket and those which do not. Our notation will sometimes reflect this.

For example, assume G has a left action on M. One can then show that the construction of fundamental vector fields defines a Lie algebra anti-homomorphism $\mathfrak{g} \to \mathcal{X}(M)$. Analogously one can show that every Ad_g is an automorphism of $\mathfrak{g}$, *i.e.* it preserves the Lie algebra structure: $Ad_g([v,w]) = [Ad_g(v), Ad_g(w)]$.

Let us now return to the general case of a Lie group acting on a manifold M. We can apply the above information on the geometry of Lie groups to develop a better understanding of the geometric aspects of the group action.

Definition A.14. Assume G acts on M. Fix $x \in M$. The *orbit of x* in M is the subset

$$\mathcal{O}_x := \{g \cdot x : g \in G\} \subseteq M.$$

Notice that $\mathcal{O}_{gx} = \mathcal{O}_x$. The *stabilizer of x* in G is the closed subgroup

$$G_x := \{g \in G : g \cdot x = x\} \subseteq G.$$

This is again a Lie group. We denote its Lie algebra $\mathfrak{g}_x$: it is a subalgebra of $\mathfrak{g}$. It is simple to check that $G_{gx} = I_g(G_x) = g \cdot G_x \cdot g^{-1}$ and that $\mathfrak{g}_{gx} = Ad_g(\mathfrak{g}_x)$.

We say that a subset $\mathcal{O} \subseteq M$ is an *orbit* of the action if $\mathcal{O} = \mathcal{O}_x$, for some $x \in M$.

The differential geometry of an orbit can be studied via the theory of *homogeneous manifolds*, *i.e.* manifolds obtained as quotients of Lie groups, as follows.

Lemma A.15. *Let G be a Lie group and H be a closed subgroup. Then:*
  (1) *The quotient space G/H has a natural smooth structure such that the projection $\pi : G \to G/H$ is a smooth map. The differential $\nabla\pi : T_e G \to T_{[e]}(G/H)$ is surjective with kernel $T_e H$ so it yields an identification $T_{[e]}(G/H) \simeq T_e G/T_e H$.*
  (2) *Left multiplication defines a natural action of G on the manifold G/H such that π is G-equivariant. Choose $v \in T_e G$. Then the corresponding fundamental vector field on G/H, evaluated at $[e]$, coincides with $\nabla\pi(v)$.*
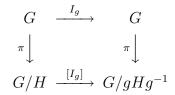
(3) *Now assume G acts on a manifold M. Choose $x \in M$ and set $H := G_x$. Then the group action defines a smooth 1:1 equivariant immersion (not necessarily an embedding)*

$$j : G/H \to M, \quad j([g]) := g \cdot x$$

with image $\mathcal{O}_x$. Using this immersion we can thus identify $\mathcal{O}_x$ with G/G_x.

Lemma A.15 identifies the geometry of $\mathcal{O}_x$ with the geometry of the homogeneous space G/G_x. The choice of point x plays an important role in this identification. However, given an orbit $\mathcal{O}$ of G in M, the choice of $x \in \mathcal{O}$ is not canonical. Furthermore, given any other point $y \in \mathcal{O}$, the choice of $g \in G$ such that $y = gx$ is also not unique. The following lemma shows how things change under different such choices.

Lemma A.16. *Let G be a Lie group and H be a closed subgroup. Choose $g \in G$ and let I_g denote the corresponding inner automorphism of G. It is simple to check that setting $[I_g]([k]) := [I_g(k)]$ yields a well-defined commutative diagram*

$$
\begin{array}{ccc}
G & \xrightarrow{I_g} & G \\
\pi \downarrow & & \pi \downarrow \\
G/H & \xrightarrow{[I_g]} & G/gHg^{-1}
\end{array}
$$

Now assume G acts on M. Choose an orbit $\mathcal{O}$ and points $x, gx \in \mathcal{O}$. Set $H := G_x$ so that we can identify $G/H \simeq \mathcal{O}_x$, $G/gHg^{-1} \simeq \mathcal{O}_{gx}$. In terms of these identifications, the map $[I_g]$ corresponds to the map $g : \mathcal{O}_x \to \mathcal{O}_{gx}$. Taking the differential of the maps in the above diagram thus leads to the commutative diagram

$$
\begin{array}{ccc}
T_e G & \xrightarrow{Ad_g} & T_e G \\
\downarrow & & \downarrow \\
T_x \mathcal{O} & \xrightarrow{\nabla g} & T_{gx} \mathcal{O}
\end{array}
$$

where the vertical maps are those defined by the construction of fundamental vector fields.

Proof: Choose any $k \in G$. The identification $j : G/G_x \simeq \mathcal{O}_x$ implies $[k] \simeq k \cdot x$. Using the analogous identifications for gx we find

$$[I_g]([k]) = [gkg^{-1}] \simeq gkg^{-1} \cdot gx = g(k \cdot x).$$

This proves that under these identifications $[I_g]$ corresponds to g. Now choose $v = d/dt(g_t)_{|t=0} \in T_e G$. Then, using the identification $G/G_x \simeq \mathcal{O}_x$,

$$\nabla \pi(v) = d/dt(\pi(g_t))_{|t=0} = d/dt([g_t])_{|t=0} = d/dt([g_t \cdot e])_{|t=0} = d/dt(g_t \cdot [e])_{|t=0}$$
$$\simeq d/dt(g_t \cdot x)_{|t=0}.$$

This proves that $\nabla \pi(v)$ corresponds to $\tilde{v}(x)$, where $\tilde{v}$ is the fundamental vector field on $\mathcal{O}_x$ defined by v. QED.

A.3. **Cohomology and invariant cohomology.** Let M be a manifold. Recall that the *de Rham cohomology* groups of M are defined as the quotient spaces

$$H^k(M;\mathbb{R}) := \frac{\text{Ker}(d : \Lambda^k(M) \to \Lambda^{k+1}(M))}{\text{Im}(d : \Lambda^{k-1}(M) \to \Lambda^k(M))}.$$

Given an action of a group G on M, one can restrict one's attention to the space of k-forms which are invariant under the action of G. An analogous construction then leads to the definition of the *invariant de Rham cohomology* groups of the pair (M, G), cf. *e.g.* [11] Section V.12. For our purposes it is sufficient to only consider pairs of the form (G, H), where H is a closed subgroup of G acting via right multiplication, *i.e.* $R_h(g) := gh$. The construction is then as follows.

Consider the space of H-invariant k-forms on G,

$$\Lambda^k(G^H) := \{\alpha \in \Lambda^k(G) : R_h^*\alpha = \alpha, \ \ \forall h \in H\}.$$

Notice that the standard exterior differentiation operator d on $\Lambda^k(G)$ is H-equivariant, *i.e.* $d(R_h^*\alpha) = R_h^*(d\alpha)$. It thus restricts to an operator

$$d : \Lambda^k(G^H) \to \Lambda^{k+1}(G^H),$$

defining the invariant cohomology groups

$$H^k(G^H;\mathbb{R}) := \frac{\text{Ker}(d : \Lambda^k(G^H) \to \Lambda^{k+1}(G^H))}{\text{Im}(d : \Lambda^{k-1}(G^H) \to \Lambda^k(G^H))}.$$

There is a natural relationship between the invariant cohomology of the pair (G, H) and the cohomology of the manifold G/H, as follows. The projection $\pi : G \to G/H$ satisfies $\pi \circ R_h = \pi$. This implies that the pull-back operation induces injections

(A.18) $$\pi^* : \Lambda^k(G/H) \to \Lambda^k(G^H).$$

Since π^* commutes with d it defines homomorphisms between the corresponding cohomology groups

$$\pi^* : H^k(G/H;\mathbb{R}) \to H^k(G^H;\mathbb{R}), \ \ \pi^*[\alpha] := [\pi^*\alpha].$$

In the special case $k = 1$, this map is an injection. Indeed, given $[\alpha] \in H^1(G/H;\mathbb{R})$, assume $\pi^*[\alpha] = 0$. Then $\pi^*\alpha \in \Lambda^1(G^H)$ is exact, *i.e.* $\pi^*\alpha = df$ for some $f \in \Lambda^0(G/H)$. However it is clear that Equation A.18 is an isomorphism for $k = 0$, *i.e.* $f = \pi^*f'$ for some $f' \in \Lambda^0(G/H)$. Thus $\pi^*(\alpha - df') = 0$ so $\alpha = df'$, *i.e.* $[\alpha] = 0$.

Now assume given a left action of G on a manifold M. Choose an orbit $\mathcal{O}$ of this action. According to Lemma A.15, $\mathcal{O}$ is a smooth submanifold of M. Choosing $x \in \mathcal{O}$ allows us to define the invariant cohomology of the pair (G, G_x). Using the point $y = gx$ leads us instead to the invariant cohomology of the pair (G, G_y). We can use Lemma A.16 to build isomorphisms between these groups. In this sense, these cohomology groups depend only on $\mathcal{O}$. It thus makes sense to look for a construction of cohomology groups which is independent of the choice of point. This can be done as follows.

Consider the set of smooth maps from $\mathcal{O}$ into the vector space $\Lambda^k(\mathfrak{g})$,

$$\Lambda^k(\mathcal{O};\mathfrak{g}) := C^\infty(\mathcal{O}, \Lambda^k(\mathfrak{g})).$$

Notice that, given $\alpha \in \Lambda^k(\mathcal{O};\mathfrak{g})$ and $v \in \mathfrak{g}$, we obtain by contraction an element

$$i_v\alpha := \alpha(v, \cdot, \ldots, \cdot) \in \Lambda^{k-1}(\mathcal{O};\mathfrak{g}).$$

Likewise, given $v_1, \ldots, v_k \in \mathfrak{g}$, iterated contractions define an element

$$\alpha(v_1, \ldots, v_k) \in \Lambda^0(\mathcal{O}; \mathfrak{g}) = C^\infty(\mathcal{O}, \mathbb{R}).$$

For any given $v_0, \ldots, v_k \in \mathfrak{g}$ it thus makes sense to define

$$(A.19) \qquad d\alpha(v_0, \ldots, v_k) := \sum_{j=0}^{k} (-1)^j \tilde{v}_j \alpha(v_0, \ldots, \hat{v}_j, \ldots, v_k)$$
$$+ \sum_{j<l} (-1)^{j+l} \alpha(-[v_j, v_l], v_0, \ldots, \hat{v}_j, \ldots, \hat{v}_l, \ldots, v_k)$$

where $[\cdot, \cdot]$ denotes the Lie bracket on $\mathfrak{g}$, the superscript $\hat{}$ denotes an omitted term and $\tilde{v}_i$ denotes the fundamental vector field associated to v_i. One can check (or it follows from Proposition A.17) that $d\alpha \in \Lambda^{k+1}(\mathcal{O}; \mathfrak{g})$ and that $d(d\alpha) = 0$. We thus obtain cohomology groups

$$H^k(\mathcal{O}; \mathfrak{g}) := \frac{\mathrm{Ker}(d : \Lambda^k(\mathcal{O}; \mathfrak{g}) \to \Lambda^{k+1}(\mathcal{O}; \mathfrak{g}))}{\mathrm{Im}(d : \Lambda^{k-1}(\mathcal{O}; \mathfrak{g}) \to \Lambda^k(\mathcal{O}; \mathfrak{g}))}.$$

Now recall that, for any $x \in \mathcal{O}$, Lemma A.15 defines a projection $\mathfrak{g} \to T_x\mathcal{O}$. Dually, this implies that there exist natural injections $\Lambda^k(T_x\mathcal{O}) \to \Lambda^k(\mathfrak{g})$. We can use these to define injections

$$(A.20) \qquad \Lambda^k(\mathcal{O}) \to \Lambda^k(\mathcal{O}; \mathfrak{g}), \quad \Lambda \mapsto \bar{\Lambda}.$$

Proposition A.17. *Given any $x \in \mathcal{O}$ and setting $H := G_x$, there exists a canonical isomorphism $\Lambda^k(\mathcal{O}; \mathfrak{g}) \to \Lambda^k(G^H)$ leading to a commutative diagram*

$$\begin{array}{ccc} \Lambda^k(\mathcal{O}) & \xrightarrow{j^*} & \Lambda^k(G/H) \\ \downarrow & & \downarrow{\pi^*} \\ \Lambda^k(\mathcal{O}; \mathfrak{g}) & \longrightarrow & \Lambda^k(G^H) \end{array}$$

where the vertical arrow on the left denotes the map of Equation A.20. This isomorphism also leads to a canonical isomorphism between the corresponding cohomology groups, i.e. an isomorphism $H^k(\mathcal{O}; \mathfrak{g}) \to H^k(G^H; \mathbb{R})$. In particular, $H^1(\mathcal{O}; \mathbb{R})$ can be canonically viewed as a subgroup of $H^1(\mathcal{O}; \mathfrak{g})$.

Proof: Fundamental vector fields provide an identification $T_eG \to T_gG$ for any $g \in G$, *i.e.* a *parallelization* of G. Using this parallelization we can identify the space $\Lambda^k(G)$ of all k-forms on G with the space of smooth maps $G \to \Lambda^k(T_eG)$. Restricting this identification gives identifications

$$\Lambda^k(G^H) \simeq \{\alpha : G \to \Lambda^k(T_eG) : \alpha(gh) = \alpha(g), \ \forall h \in H\}$$
$$\simeq C^\infty(G/H, \Lambda^k(T_eG))$$
$$= C^\infty(\mathcal{O}, \Lambda^k(\mathfrak{g}))$$
$$= \Lambda^k(\mathcal{O}; \mathfrak{g}).$$

It is simple to check that, up to these identifications, the above diagram commutes.

Now choose $\alpha \in \Lambda^k(\mathcal{O}; \mathfrak{g})$. Let α' denote the corresponding element of $\Lambda^k(G^H)$. As usual let us denote by $\tilde{v}$ the fundamental vector field generated by v. We now want to prove that $d(\alpha') = (d\alpha)'$, *i.e.* that, for all $v_0, \ldots, v_k \in T_e G$,

$$(A.21) \qquad\qquad d(\alpha')(\tilde{v}_0, \ldots, \tilde{v}_k) = d\alpha(v_0, \ldots, v_k).$$

According to Equation A.11, we can calculate the left hand side using the usual bracket on $\mathcal{X}(G)$. However, recall from Remark A.11 that $[\tilde{v}_i, \tilde{v}_j] = -\widetilde{[v_i, v_j]}$. The change of sign here is cancelled by the choice of signs in Equation A.19. This proves the claim on d, thus on the cohomology groups.

Clearly there also exists an identification $j^* : H^1(\mathcal{O}; \mathbb{R}) \simeq H^1(G/H; \mathbb{R})$. We can now use the injection $\pi^* : H^1(G/H; \mathbb{R}) \to H^1(G^H; \mathbb{R})$ to prove the last claim. QED.

A.4. **The group of diffeomorphisms.** Let $\mathrm{Diff}_c(\mathbb{R}^D)$ denote the set of diffeomorphisms of $\mathbb{R}^D$ with compact support, *i.e.* those which coincide with the identity map Id outside of a compact subset of $\mathbb{R}^D$. Composition of maps clearly yields a group structure on $\mathrm{Diff}_c(\mathbb{R}^D)$. It is possible to endow $\mathrm{Diff}_c(\mathbb{R}^D)$ with the structure of an infinite-dimensional Lie group in the sense of [37]. A local model is provided by the space $\mathcal{X}_c(\mathbb{R}^D)$, endowed as in Section 2.1 with the structure of a topological vector space. More specifically, we can apply the construction outlined in Remark A.21 below to build a local chart $\mathcal{U}$ for $\mathrm{Diff}_c(\mathbb{R}^D)$ near the identity element Id. This yields by definition an isomorphism $T_{Id}\mathrm{Diff}_c(\mathbb{R}^D) \simeq \mathcal{X}_c(\mathbb{R}^D)$. We can then use right multiplication to build charts $\mathcal{U}_\phi := \{u \circ \phi : u \in \mathcal{U}\}$ around any $\phi \in \mathrm{Diff}_c(\mathbb{R}^D)$, leading to $T_\phi \mathrm{Diff}_c(\mathbb{R}^D) \simeq \{X \circ \phi : X \in \mathcal{X}_c(\mathbb{R}^D)\}$. Thoughout this article we will generally restrict our attention to the connected component of $\mathrm{Diff}_c(\mathbb{R}^D)$ containing Id.

Remark A.18. It may be useful to emphasize that defining charts on $\mathrm{Diff}_c(\mathbb{R}^D)$ as above leads to the following interpretation of Equation A.8: ϕ_t is a smooth path on $\mathrm{Diff}_c(\mathbb{R}^D)$ and $X_t \circ \phi_t \in T_{\phi_t}\mathrm{Diff}_c(\mathbb{R}^D)$ is its tangent vector field.

As usual one can define the Lie algebra to be the tangent space at Id. The Lie bracket $[\cdot, \cdot]_{\mathfrak{g}}$ on this space can then be defined as in Section A.2, cf. [37].

Lemma A.19. *The adjoint representation of $\mathrm{Diff}_c(\mathbb{R}^D)$ on $\mathcal{X}_c(\mathbb{R}^D)$ coincides with the push-forward operation: $Ad_\phi(X) = \phi_*(X)$. Furthermore, the Lie bracket on $\mathcal{X}_c(\mathbb{R}^D)$ induced by the Lie group structure on $\mathrm{Diff}_c(\mathbb{R}^D)$ is the negative of the standard Lie bracket on vector fields.*

Proof: Assume that X integrates to $\phi_t \in \mathrm{Diff}_c(\mathbb{R}^D)$. Then

$$Ad_\phi(X) = d/dt(\phi \circ \phi_t \circ \phi^{-1})_{|t=0} = \nabla\phi_{|\phi^{-1}} \cdot X_{|\phi^{-1}} = \phi_*(X).$$

As in Lemma A.12 we can calculate the Lie bracket by differentiating the adjoint representation. Thus:

$$[X, Y]_{\mathfrak{g}} = d/dt(Ad_{\phi_t}Y)_{|t=0} = d/dt(\phi_{t*}Y)_{|t=0} = -[X, Y].$$

QED.

Remark A.20. Lemma A.19 explains why the map of Remark A.13 is an algebra anti-homomorphism. Indeed, any left group action defines a homomorphism $G \to \mathrm{Diff}(M)$, thus a homomorphism between the corresponding Lie algebras. However, we now see that

the bracket used for $\mathcal{X}(M)$ in Remark A.13 is the negative of the bracket induced by the Lie group structure. Lemma A.19 is also related to Remark A.11.

Remark A.21. A similar construction proves that for any compact (respectively, noncompact) manifold M the group of diffeomorphisms $\mathrm{Diff}(M)$ (respectively, $\mathrm{Diff}_c(M)$) is an infinite-dimensional Lie group in the sense of [37]. Some care has to be exercised however in all these constructions, specifically in the definition of the local chart near Id. The naive choice

$$\mathcal{X}(M) \to \mathrm{Diff}(M), \quad X \mapsto \phi_1,$$

where ϕ_1 is the time $t = 1$ diffeomorphism obtained by integrating X to the flow ϕ_t, is not possible as it does not cover an open neighbourhood of Id, cf. [37] Warning 1.6. Instead, the standard trick is to notice that diffeomorphisms near Id are in a 1:1 relationship (via their graphs) with smooth submanifolds close to the diagonal $\Delta \subset M \times M$. These submanifolds can then be parametrized as follows. Assume $E \to M$ is a vector bundle over M. Let Z denote its zero section and U denote an open neighbourhood of Z. Assume one can find a diffeomorphism $\zeta : U \to M \times M$ sending Z to Δ. Then diffeomorphisms of M near Id correspond to smooth submanifolds of E near Z, *i.e.* smooth sections. For example, to construct a chart for diffeomorphisms close to Id we would use $E := TM$ setting ζ to be the Riemannian exponential map (with respect to a fixed metric on M).

Good choices of E and ζ for $\mathrm{Diff}(M)$ can yield as a by-product the fact that specific subgroups G of $\mathrm{Diff}(M)$ also admit Lie group structures such that the natural immersion $G \to \mathrm{Diff}(M)$ is smooth. For example, to prove this fact for the subgroups of symplectomorphisms or Hamiltonian diffeomorphisms of a symplectic manifold (M, ω) (see Section 6.1) one can choose $E := T^*M$ and the ζ defined by Weinstein's "Lagrangian neighbourhood theorem", cf. [46] Section 6 or [36] Proposition 3.34.

References

1. M. Agueh, N. Ghoussoub, and X. Kang, *Geometric inequalities via a general comparison principle for interacting gases*, Geom. Funct. Anal. **14** (2004), no. 1, 215–244.
2. Martial Agueh, *Asymptotic behavior for doubly degenerate parabolic equations*, C. R. Math. Acad. Sci. Paris **337** (2003), no. 5, 331–336.
3. Luigi Ambrosio and Wilfred Gangbo, *Hamiltonian ODEs in the Wasserstein space of probability measures*, Comm. Pure Appl. Math. **61** (2008), no. 1, 18–53.
4. Luigi Ambrosio and Nicola Gigli, *Construction of the parallel transport in the wasserstein space*, Methods Appl. Anal. (to appear).

5. Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2005.

6. P. L. Antonelli, D. Burghelea, and P. J. Kahn, *The non-finite homotopy type of some diffeomorphism groups*, Topology **11** (1972), 1–49.

7. V. Arnold, *Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits*, Ann. Inst. Fourier (Grenoble) **16** (1966), no. fasc. 1, 319–361.

8. Iwo Białynicki-Birula, John C. Hubbard, and Łukasz A. Turski, *Gauge-independent canonical formulation of relativistic plasma theory*, Phys. A **128** (1984), no. 3, 509–519.

9. M. Born, *On the quantum theory of the electromagnetic field*, Proceedings of the Royal Society of London. Series A **143** (1934), no. 849, 410–437.

10. M. Born and L. Infeld, *On the quantization of the new field theory. ii*, Proceedings of the Royal Society of London. Series A **150** (1935), no. 869, 141–166.

11. Glen E. Bredon, *Topology and geometry*, Graduate Texts in Mathematics, vol. 139, Springer-Verlag, New York, 1993.

12. E. A. Carlen and W. Gangbo, *Constrained steepest descent in the 2-Wasserstein metric*, Ann. of Math. (2) **157** (2003), no. 3, 807–846.

13. José A. Carrillo, Robert J. McCann, and Cédric Villani, *Contractions in the 2-Wasserstein length space and thermalization of granular media*, Arch. Ration. Mech. Anal. **179** (2006), no. 2, 217–263.

14. Hernán Cendra, Darryl D. Holm, Mark J. W. Hoyle, and Jerrold E. Marsden, *The Maxwell-Vlasov equations in Euler-Poincaré form*, J. Math. Phys. **39** (1998), no. 6, 3138–3157.

15. Paul R. Chernoff and Jerrold E. Marsden, *Properties of infinite dimensional Hamiltonian systems*, Lecture Notes in Mathematics, Vol. 425, Springer-Verlag, Berlin, 1974.

16. D. Cordero-Erausquin, B. Nazaret, and C. Villani, *A mass-transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities*, Adv. Math. **182** (2004), no. 2, 307–332.

17. Dario Cordero-Erausquin, Wilfrid Gangbo, and Christian Houdré, *Inequalities for generalized entropy and optimal transportation*, Recent advances in the theory and applications of mass transport, Contemp. Math., vol. 353, Amer. Math. Soc., Providence, RI, 2004, pp. 73–94.

18. David G. Ebin and Jerrold Marsden, *Groups of diffeomorphisms and the notion of an incompressible fluid.*, Ann. of Math. (2) **92** (1970), 102–163.

19. Wilfrid Gangbo and Robert J. McCann, *The geometry of optimal transportation*, Acta Math. **177** (1996), no. 2, 113–161.

20. Wilfrid Gangbo and Adrian Tudorascu, in preparation.

21. Victor Guillemin and Alan Pollack, *Differential topology*, Prentice-Hall Inc., Englewood Cliffs, N.J., 1974.

22. Richard S. Hamilton, *The inverse function theorem of Nash and Moser*, Bull. Amer. Math. Soc. (N.S.) **7** (1982), no. 1, 65–222.

23. Richard Jordan, David Kinderlehrer, and Felix Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal. **29** (1998), no. 1, 1–17 (electronic).

24. Boris Khesin and Paul Lee, *Poisson geometry and first integrals of geostrophic equations*, Phys. D **237** (2008), no. 14-17, 2072–2077.

25. Hwa Kil Kim, *Ph.D. thesis, Georgia Institute of Technology*, (2009).

26. Shoshichi Kobayashi and Katsumi Nomizu, *Foundations of differential geometry. Vol I*, Interscience Publishers, a division of John Wiley & Sons, New York-London, 1963.

27. Andreas Kriegl and Peter W. Michor, *The convenient setting of global analysis*, Mathematical Surveys and Monographs, vol. 53, American Mathematical Society, Providence, RI, 1997.

28. John Lott, *Some geometric calculations on Wasserstein space*, Comm. Math. Phys. **277** (2008), no. 2, 423–437.

29. John Lott and Cédric Villani, *Ricci curvature for metric-measure space via optimal transport*, Ann. of Math. (to appear).

30. Francesco Maggi and Cédric Villani, *Balls have the worst best Sobolev inequalities*, J. Geom. Anal. **15** (2005), no. 1, 83–121.

31. J. E. Marsden, A. Weinstein, T. Ratiu, R. Schmid, and R. G. Spencer, *Hamiltonian systems with symmetry, coadjoint orbits and plasma physics*, Proceedings of the IUTAM-ISIMM symposium on modern developments in analytical mechanics, Vol. I (Torino, 1982), vol. 117, 1983, pp. 289–340.

32. Jerrold E. Marsden and Tudor S. Ratiu, *Introduction to mechanics and symmetry*, second ed., Texts in Applied Mathematics, vol. 17, Springer-Verlag, New York, 1999.
33. Jerrold E. Marsden and Alan Weinstein, *The Hamiltonian structure of the Maxwell-Vlasov equations*, Phys. D **4** (1981/82), no. 3, 394–406.
34. Robert J. McCann, *A convexity principle for interacting gases*, Adv. Math. **128** (1997), no. 1, 153–179.
35. Robert J. McCann and Peter Topping, *Ricci flow, entropy and optimal transportation*.
36. Dusa McDuff and Dietmar Salamon, *Introduction to symplectic topology*, second ed., Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 1998.
37. J. Milnor, *Remarks on infinite-dimensional Lie groups*, Relativity, groups and topology, II (Les Houches, 1983), North-Holland, Amsterdam, 1984, pp. 1007–1057.
38. Hideki Omori, *Infinite-dimensional Lie groups*, Translations of Mathematical Monographs, vol. 158, American Mathematical Society, Providence, RI, 1997, Translated from the 1979 Japanese original and revised by the author.
39. K. Ono, *Floer-Novikov cohomology and the flux conjecture*, Geom. Funct. Anal. **16** (2006), no. 5, 981–1020.
40. Felix Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations **26** (2001), no. 1-2, 101–174.
41. Wolfgang Pauli, *General principles of quantum mechanics*, Springer-Verlag, Berlin, 1980, Translated from the German by P. Achuthan and K. Venkatesan, With an introduction by Charles P. Enz.
42. Walter Rudin, *Functional analysis*, second ed., International Series in Pure and Applied Mathematics, McGraw-Hill Inc., New York, 1991.
43. Karl-Theodor Sturm, *On the geometry of metric measure spaces. I*, Acta Math. **196** (2006), no. 1, 65–131.
44. _____, *On the geometry of metric measure spaces. II*, Acta Math. **196** (2006), no. 1, 133–177.
45. Cédric Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003.
46. Alan Weinstein, *Symplectic manifolds and their Lagrangian submanifolds*, Advances in Math. **6** (1971), 329–346 (1971).

Georgia Institute of Technology, Atlanta GA, USA
E-mail address: gangbo@math.gatech.edu

Georgia Institute of Tecnology, Atlanta GA, USA
E-mail address: hwakil@math.gatech.edu

Mathematical Institute, Oxford, UK
E-mail address: pacini@maths.ox.ac.uk