Distributions that arise as derivatives of families of measures

Rodolfo Ríos-Zertuche

Abstract

We characterize the distributions that arise as derivatives of families of probabilities and of positive and signed Borel measures on smooth manifolds.

1 Introduction

Let $P$ be a $C^\infty$ manifold without boundary, and let $\mu$ be a Borel measure on $P$. Denote by $C_c^\infty(P)$ the space of smooth functions with compact support on $P$.

Consider a family of measures $\mu_s$ indexed by a real parameter $s$ with values in an interval that contains 0, and such that $\mu = \mu_0$. We say that $\mu_s$ is differentiable at 0 if for every $f \in C_c^\infty(P)$ the function $s \mapsto \int f \, d\mu_s$ is differentiable at 0 and if the derivative induces a distribution. For example, if the family is a moving Dirac delta $\mu_s = \delta_s$ on $\mathbb{R}$, then the derivative at 0 is the distribution $-\partial \delta_0$ given by $\langle -\partial \delta_0, f \rangle = f'(0)$ for all $f \in C_c^\infty(\mathbb{R})$.

We address the question of characterizing the distributions that arise in this way. In other words, we characterize the velocity vectors for curves in the space of measures that pass through $\mu$.

We find that if the measures $\mu_s$ are allowed to be signed (i.e., to have both positive and negative mass), then any distribution can arise; see Proposition 4. On the other hand, if the measures $\mu_s$ are only allowed to be positive, we find a necessary and sufficient condition for a given distribution to be the velocity vector of a curve through $\mu_0$. This is Condition 3 below, which says that the nullspace of the distribution must contain all smooth, nonnegative functions that vanish on the support of $\mu_0$. This characterization is our main result, given in Theorem 7. This theorem also accounts for the case in which all the measures $\mu_s$ are probabilities.

Interest in the variational structure of the space of measures, which we study here, comes from the applications that the analysis of measures has found for example in problems of optimal transport (e.g., 1) and optimization, as in Mather-Aubry theory (e.g., 2, 4). Differentiable families of measures have also been studied extensively for example in 6. Our own applications of the results of this paper will appear elsewhere 5.
We give precise definitions and some preliminaries in Section 2 and we state and prove our result in Section 3.

Acknowledgements The author is very grateful to Camilo Arias Abad and Jesús Puente Arrubarrena for several conversations on this topic. He is also deeply indebted to the Max Planck Institute for Mathematics in Bonn, Germany, for a postdoctoral fellowship that made this research possible.

2 Distributions and measures

2.1 Convolutions

A \( \psi \in C^\infty_c(\mathbb{R}) \) is a function such that \( \psi(x) = \psi(-x), \int \psi = 1 \), and \( \psi \geq 0 \).

We will say that a tuple of vector fields \( F = (F_1, \ldots, F_\ell) \) on \( P \) is generating if at every point \( p \in P \) the vectors \( F_1(p), \ldots, F_\ell(p) \) span all of the tangent space \( T_pP \).

Fix a generating tuple of vector fields \( F = (F_1, \ldots, F_\ell) \). Denote by \( \phi_i : P \times \mathbb{R} \to P \) the flow of \( F_i \):

\[
\phi_0^i(x) = 0, \quad \frac{d\phi_s^i(x)}{ds} = F_i(\phi_s^i(x)), \quad s \in \mathbb{R}.
\]

For \( f \in C^\infty_c(P) \), we will denote by \( P_i(f) \) the function given by

\[
P_i(f)(x) = \int_{\mathbb{R}} f \circ \phi_s^i(x) \psi(s) \, ds.
\]

This is a convolution in the direction \( F_i \).

For \( f \in C^\infty_c(P) \), we will denote

\[
\psi * F f := P_1 P_2 \cdots P_\ell(f).
\]

2.2 Definition and smoothing of distributions

A distribution on the open set \( U \subseteq \mathbb{R}^m \) is a linear functional \( \eta : C^\infty_c(U) \to \mathbb{R} \) such that for each compact set \( K \subset U \) there are some constants \( N > 0 \) and \( C > 0 \) (depending only on \( K \) and \( \eta \)) such that

\[
|\langle \eta, f \rangle| \leq C \sum_{|I| \leq N} \sup_{p \in U} |\partial^I f(p)|
\]

for all \( f \in C^\infty_c(U) \). Here, the sum is taken over all multi-indices \( I \) with \( m \) nonnegative entries adding up to at most \( N \), and \( \partial^I \) denotes the iterated partial derivatives in the corresponding directions in \( \mathbb{R}^m \).
We fix, once and for all, an $n$-dimensional $C^\infty$ manifold $P$ without boundary, and with a Riemannian metric that induces the distance $\text{dist}_P$ between points of $P$.

Let $\eta: C^\infty_c(P) \to \mathbb{R}$ be a linear functional. For a chart $\varepsilon: U \to W$ from the open set $U \subseteq P$ to the open set $W \subseteq \mathbb{R}^n$, the pushforward $\varepsilon_* \eta$ is defined by

$$\langle \varepsilon_* \eta, f \rangle = \langle \eta, f \circ \varepsilon \rangle$$

for $f$ in $C^\infty_c(W)$.

The functional $\eta$ is a distribution if for each chart $\varepsilon$ as above, $\varepsilon_* \eta$ is a distribution on $W$. We will denote by $\mathcal{D}'(P)$ the space of distributions on $P$. The topology on $\mathcal{D}'(P)$ is induced by the seminorms

$$\eta \mapsto |\langle \eta, f \rangle|$$

for $f \in C^\infty_c(P)$. In other words, we have $\eta_i \to \eta$ if, and only if, $\langle \eta_i, f \rangle \to \langle \eta, f \rangle$ for all $f \in C^\infty_c(P)$. We remark that any measure on $P$ determines a distribution, but that not all distributions arise in this way.

For a distribution $\eta \in \mathcal{D}'(P)$, we define the convolution by duality:

$$\langle \psi *_F \eta, f \rangle = \langle \eta, \psi *_F f \rangle.$$  

**Lemma 1.** If $\eta$ is a distribution in $\mathcal{D}'(P)$, $F$ is a generating tuple of vector fields, and $\psi$ is a mollifier, then $\psi *_F \eta$ is a smooth signed Borel measure.

For a proof see for example [3, §5.2].

### 2.3 Structure

We fix a generating tuple $F = (F_1, \ldots, F_\ell)$ of vector fields. As before, we denote by $I$ a multi-index $I = (i_1, \ldots, i_\ell)$ with $\ell$ nonnegative entries, and by $\partial^I$ the operator that iteratively takes $i_j$ covariant derivatives in the direction $F_j$, $j = 1, \ldots, \ell$.

As usual in the theory of distributions, we define derivatives of distributions $\nu$ by duality,

$$\langle \partial^I \nu, f \rangle = (-1)^{|I|} \langle \nu, \partial^I f \rangle,$$

and the support $\text{supp} \nu$ of a distribution $\nu$ to be largest set such that if $f \in C^\infty_c(P)$ is supported outside $\text{supp} \nu$ then $\langle \nu, f \rangle = 0$.

**Lemma 2** (Structural representation in terms of measures). A distribution $\eta \in \mathcal{D}'(P)$ can be written as a sum

$$\eta = \sum_I \partial^I \nu_I$$  

where $I$ ranges over all multi-indices as above; for each $I$, $\nu_I$ is a signed measure. For a compact set $K \subseteq P$,

$$K \cap \text{supp} \nu_I = \emptyset$$
for all but finitely many multi-indices $I$.

Proof. Take a partition of unity $\{\xi_j\}_{j \in \mathbb{N}} \subseteq C_0^\infty(P)$ of $P$, that is, a countable set of smooth functions $\xi_i$ with compact support such that $\sum_j \xi_j(p) = 1$ and, on each compact set $K \subseteq P$, the restriction $\xi_j|_K \equiv 0$ for all but finitely many $j \in \mathbb{N}$. We make the further assumption that the support of each of the functions $\xi_j$ is contained in an open set $U_j \subseteq P$ that is diffeomorphic to a cube $(0,1)^n$, and we let $\phi_j : U_j \to (0,1)^n$ be the corresponding diffeomorphism.

We let $\tilde{\eta}_j$ be the distribution on $\mathbb{R}^n$ that results from pushing $\xi_j \eta$ forward to the cube $(0,1)^n$ and extending periodically. In other words, for all rapidly-decreasing (Schwartz) functions $f \in C_0^\infty(\mathbb{R}^n)$, we let $\tau_z f(x) = f(x-z)$ and

$$\langle \tilde{\eta}_j, f \rangle = \sum_{z \in \mathbb{Z}^n} \langle \eta, \xi_j \cdot (\tau_z f) \circ \phi_j \rangle.$$

Like all periodic distributions, $\tilde{\eta}_j$ is a tempered distribution. We have

**Lemma 3.** Every tempered distribution is a derivative of finite order of some continuous function of polynomial growth.

For a proof, see for example [3, Theorem 3.8.1].

Let $\zeta_j$ be the continuous function of polynomial growth corresponding to $\tilde{\eta}_j$ (as furnished by Lemma 3) and let $I_j$ be the multi-index corresponding to the derivative in the lemma, so that

$$\tilde{\eta}_j = \partial^{I_j} \zeta_j.$$

Let $D_j$ be the (smooth) differential operator on $P$ such that $\phi_j^* \partial^{I_j} = D_j \phi_j^*$, where $\phi_j^*$ denotes the pullback by $\phi_j$. Observe that

$$\xi_j \eta = \phi_j^* \partial^{I_j} \zeta_j = D_j \phi_j^* \zeta_j.$$

Since $\zeta_j$ is a continuous function, $\phi_j^* \zeta_j$ is piecewise continuous, and hence it induces a measure on $U_j$. Then we can write

$$\eta = \sum_j \eta_j \eta = \sum_j D_j \phi_j^* \zeta_j,$$

and since each of the summands on the right can be expressed as a finite sum of derivatives of a continuous function, this proves the lemma.

3 Variations

Let $\mu_s$ be a family of Borel measures on the manifold $P$ parameterized by a real parameter $s$ with values in an open interval $J \subseteq \mathbb{R}$ that contains 0.
We say that the family $\mu_s$ is differentiable at 0 if there is a distribution $\eta \in \mathcal{D}'(P)$ such that, for every function $f \in C_\infty^c(P)$,

$$\frac{d}{ds}|_{s=0} \int f \, d\mu_s = \langle \eta, f \rangle.$$

The distribution $\eta$ is the derivative $d\mu_s/|_{s=0}$ of $\mu_s$ at 0.

**Remark 4.** This is just one way to define differentiability of families of distributions; other ways have been explored for example in [6].

**Proposition 5.** For every Borel measure $\mu$ and every distribution $\eta$, there exists a family of signed measures with

$$\mu_0 = \mu \quad \text{and} \quad \frac{d\mu_s}{ds}|_{s=0} = \eta.$$

We prove this below. For the proof, we need to define a family of distributions $\eta_s$ to be differentiable if there is a distribution $\nu$ such that for every function $f \in C_\infty^c(P)$,

$$\frac{d}{ds}|_{s=0} \langle \eta_s, f \rangle = \langle \nu, f \rangle.$$

**Lemma 6.** For a generating tuple $F$ of vector fields, a mollifier $\psi$, and any family of distributions $\eta_s$ differentiable at $s = 0$, we have

$$\frac{d}{ds}|_{s=0} \psi *_{sF} \eta_s = \frac{d\eta_s}{ds}|_{s=0}.$$

**Proof.** This follows from the fact that for $f \in C_\infty^c(P)$, $s \mapsto \psi *_{sF} f$ is an even function, so its derivative at $s = 0$ must vanish.

**Proof of Proposition 5.** Take a mollifier $\psi$ and a generating tuple $F$ of vector fields. Then, as follows from Lemmas 6 and 6, the family $\mu_s = \psi *_{sF} (\mu_0 + s\nu)$ has the required properties.

For families of positive measures, the situation is different.

**Theorem 7.** Let $\mu$ be a positive Borel measure and let $\eta$ be a distribution. Denote by $\mathcal{F}_\mu$ the space of nonnegative functions $f \in C_\infty^c(P)$ that vanish identically on $\text{supp} \mu$. Then there exists a family $\mu_s$ of positive measures with $\mu_0 = \mu$ and derivative $d\mu_s/|_{s=0} = \eta$ if, and only if, $\eta$ satisfies the following condition:

$$\langle \eta, f \rangle = 0 \text{ for every } f \in \mathcal{F}_\mu. \quad (D)$$

If $\mu$ is a probability measure and $\eta$ additionally satisfies that $\langle \eta, 1 \rangle = 0$, then $\mu_s$ can be realized as a family of probability measures.
Remark 8. Condition D implies that supp $\eta \subseteq$ supp $\mu$. Apart from this, Condition D is relevant only when supp $\mu$ has parts that are very thin — only one point thick.

For example, if $P = \mathbb{R}$, $\mu$ is the Dirac delta $\delta_0$, then Condition D implies that $\eta$ must be of the form $A\delta_0 + B\partial\delta_0$, $A, B \in \mathbb{R}$. Indeed, take a cutoff function $\rho: \mathbb{R} \to \mathbb{R} \in C^\infty_c(\mathbb{R})$ (i.e., $\rho \geq 0$, $\rho \equiv 1$ in a neighborhood of 0 and $\rho \equiv 0$ outside a slightly larger neighborhood), then taking $f(x) = \rho(x) \sum_{i \geq 2} c_i x^i$ (with $c_2$ large enough to ensure that $f \geq 0$) we see that $\eta$ must be of the proposed form in order to comply with the condition.

On the other hand, if we again had $P = \mathbb{R}$, but now $\mu = \chi_{[0,1]}$, the characteristic function on the unit interval, then as long as supp $\eta \subseteq$ supp $\mu$, $\eta$ can be any distribution and still comply with Condition D.

Remark 9. The family $\mu_s$ can always be realized as a family of smooth measures (except maybe at $s = 0$). Indeed, if $\mu_s$ is any family of measures that is differentiable at $s = 0$, $\psi$ is a mollifier, and $F$ is a generating tuple, then the measure $\tilde{\mu}_s = \psi \ast_s F \mu_s$ has the same derivative at 0 and the same mass as $\mu_s$, and $\tilde{\mu}_s$ is a positive measure if $\tilde{\mu}_s$ is. By Lemma 10 the measure $\tilde{\mu}_s$ is a smooth density for all $s \neq 0$.

Lemma 10. Fix a point $p \in \text{supp } \mu \subseteq P$. Let $\eta_p$ be a distribution supported on $p$ that satisfies Condition D. Then there is a family of positive measures $\mu_p^s$ such that $\mu_p^0 = \mu$ and

$$\frac{d\mu_p^s}{ds} \bigg|_{s=0} = \eta_p.$$  

Moreover, the dependence of $\mu_p^s$ on $p$ is measurable.

If $\mu$ is a probability measure and additionally $\langle \eta_p, 1 \rangle = 0$, then $\mu_p^s$ can be realized as a family of probability measures.

For the proof of the lemma we will need a metric defined on the space of distributions involving up to $k$th derivatives, $k \geq 1$, and given by

$$\text{dist}_k(\theta_1, \theta_2) = \sum_{j=1}^{\infty} \frac{1}{2^j \|f_j\|_k} |\langle \theta_1, f_j \rangle - \langle \theta_2, f_j \rangle|$$

for two distributions $\theta_1$ and $\theta_2$, and with $\{f_j\} \subset C^\infty_c(P)$ a sequence of functions that is dense with respect to the norm

$$\|f\|_k = \sum_{|I| \leq k} \sup_{q \in P} |\partial^I f(q)|.$$  

Proof of Lemma 10. Let $V \subseteq T_p P$ be the subspace that is null for the Hessians at $p$ of all the functions in $\mathcal{F}_\mu$:

$$V = \{ v \in T_p P : \text{Hess}_p f(v, v) = 0 \text{ for all } f \in \mathcal{F}_\mu \}.$$
Let \( m = \dim V \leq n = \dim P \). Take coordinates \((x_1, x_2, \ldots, x_n)\) around \( p \) such that the vectors
\[
\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_m} \in T_p P
\]
form a basis of \( V \) and \( \partial/\partial x_1, \ldots, \partial/\partial x_n \) is an orthonormal basis of \( T_p P \). Then by Lemma 2 we know that \( \eta \) must be a finite linear combination of distributions of the form
\[
\left( \frac{\partial}{\partial x_u} \right)^{e_0} \left( \frac{\partial}{\partial x_1} \right)^{e_1} \left( \frac{\partial}{\partial x_2} \right)^{e_2} \cdots \left( \frac{\partial}{\partial x_m} \right)^{e_m} \delta_p
\]
where \( e_0 \in \{0, 1\} \), \( u > m \), and the integers \( e_1, \ldots, e_m \) are nonnegative. For reasons analogous to those explained in Remark 8, Condition D makes it impossible to have higher derivatives in the directions outside \( V \).

Let us show that it is enough to improve the lemma for the case in which \( e_0 = 0 \). Indeed, if \( \nu_s \) is a family of positive measures such that \( \nu_0 = \mu \) and
\[
\frac{d\nu_s}{ds} \bigg|_{s=0} = \left( \frac{\partial}{\partial x_u} \right)^{e_1} \left( \frac{\partial}{\partial x_2} \right)^{e_2} \cdots \left( \frac{\partial}{\partial x_m} \right)^{e_m} \delta_p,
\]
and if \( \phi \) is the flow of the vector field \( \partial/\partial x_u \), then
\[
\frac{d}{ds} \phi^s \nu_s \bigg|_{s=0} = \frac{\partial}{\partial x_u} \left( \frac{\partial}{\partial x_1} \right)^{e_1} \left( \frac{\partial}{\partial x_2} \right)^{e_2} \cdots \left( \frac{\partial}{\partial x_m} \right)^{e_m} \delta_p.
\]
So we will assume that \( e_0 = 0 \) and we will focus on finding such a family \( \nu_s \). In particular, we will assume that \( \eta_p \) is of the form given in the right-hand-side of equation (2). In other words, we will assume that it only involves derivatives in the directions of \( V \).

**Lemma 11.** Let \( F \) be a generating tuple of vector fields. For \( \eta_p \) as in the right-hand-side of equation (2) and for \( k = \sum_{i=1}^m e_i \), we have, for all \( t > 0 \),
\[
\inf_g \text{dist}_k(g \cdot (\psi \ast t F \mu), \eta_p) = 0,
\]
where the infimum is taken over all measurable functions \( g: P \to \mathbb{R} \).

The reader will find the proof of Lemma 11 below.

With \( F \) as in the lemma, let
\[
r_{i,m} = \sum_{j=1}^{\infty} \frac{1}{2^{jm}} \quad \text{and} \quad \mu_i = \psi \ast r_{i,2} F \mu
\]
for \( i \in \mathbb{N} \). In particular \( r_{i,1} \to 0, r_{i,2} \to 0 \), and \( \mu_i \to \mu \) as \( i \to +\infty \). Let \( k \) be as in Lemma 11 and denote by \( \| \cdot \|_\infty \) de essential supremum norm. For each \( j \in \mathbb{N} \), take a measurable function \( g_j \) such that \( \|g_j\|_\infty \leq 1 \) and
\[
\text{dist}_k(2^j g_j \mu_j, \eta_p) < \frac{1}{2^j} + \inf_{\|g\|_\infty \leq 1} \text{dist}_k(2^j g \mu_j, \eta_p),
\]

7
where the infimum is taken over all measurable functions $g : P \to \mathbb{R}$ with essential supremum $\leq 1$. With this definition, Lemma 11 implies that if we let $j \to +\infty$, we get $2^j g_j \mu_j \to \eta_p$.

We let, for $r_{i+1,2} \leq |s| < r_{i,2}$,

$$\nu_s = (1 - 2^{1-i}) \mu_i + \sum_{j=i}^{\infty} \frac{1}{2^j} (1 + g_j \text{sgn } s) \mu_j.$$  

By construction, $\nu_s$ is a family of positive measures such that $\nu_s \to \mu$ as $s \to 0$ and its derivative at $s = 0$ is $\eta_p$. To see why, first note that, as $s \to 0,$

$$1 - 2^{1-i} \mu_i + \sum_{j=i}^{\infty} \frac{\mu_j}{2^j} \to \mu,$$

and the derivative of that term at $s = 0$ vanishes by Lemma 6. The other term vanishes as $s \to 0$, and its derivative is the limit, as $s \to 0,$ of

$$\frac{1}{s} \sum_{j=i}^{\infty} \frac{\text{sgn } s}{2^j} g_j \mu_j \approx \frac{1}{r_{i,2}} \sum_{j=i}^{\infty} \frac{\mu_j}{2^j} \to \frac{2^{-i} g_i \mu_i}{2^{-2i}} \to \eta_p,$$

where we applied L'Hôpital’s rule because both the sum and $r_{i,2}$ tend to 0 as $s \to 0$.

To ensure the measurability of the dependence of this construction in $p$, we further specify the construction as follows. For each $j \in \mathbb{Z}_-$, we take a covering of $P$ by measurable sets $A_j$ of diameter at most $1/j$. For all $p \in A_j$, we take the same function $g_j$. This ensures that these choices are made in a ‘measurable’ way. The rest of the construction does not depend on arbitrary choices, so the dependence becomes measurable.

The last statement of the lemma follows from the fact that if $\eta_p$ satisfies $\langle \eta_p, 1 \rangle = 0$, then either $g_j$ can be chosen so that $g_j \mu_j$ satisfies this too, or else $e_1 = e_2 = \cdots = e_n = 0$, and in both cases the coordinates can be picked so that the mass is preserved by the flow $\phi_s$ for small-enough $|s|$.

Proof of Lemma 11. This is a local problem and by pushing forward with a chart, we may assume that $P$ is some Euclidean space $\mathbb{R}^n$. Let $\bar{\mu} = \psi \ast_{tF} \mu$. Note that supp $\bar{\mu}$ has nonempty interior fact, and in fact contains a neighborhood of $p$.

For each $j = 1, 2, \ldots,$ let $\{x_i^j\}_{i=1}^{\infty} \subseteq \text{supp } \bar{\mu} \subset \mathbb{R}^n$ be a sequence of points contained within distance $1/j$ of $p$. We also assume that their Zariski closure is all of $\mathbb{R}^n$ (i.e., that no nonzero polynomial vanishes on all of them simultaneously). For a large-enough finite subset $I_j$ of $\mathbb{N}$, there is always a solution to the problem of finding real numbers $c_{ij}$ such that

$$\langle \eta_p, f \rangle = \lim_{h \to 0} \frac{1}{hk} \sum_{i \in I_j} c_{ij} f(hx_i^j)$$  \hspace{1cm} (3)

8
for all \( f \in C_c^\infty(P) \). To see this, note that expanding the right-hand-side as Taylor series in \( h \) and comparing coefficients, one obtains a linear system in the variables \( c_{ij} \), and that this system has solutions if sufficiently many points \( x_i^j \) are available and they are in a sufficiently-general position (which is true since their Zariski closure can be made as large as necessary). Equivalently, we have a measure that approximates \( \eta_p \):

\[
h^{-k} \sum_{i \in I_j} c_{ij} \delta_{x_i^j} \rightarrow \eta_p \quad \text{as} \quad h \to 0.
\]

These measures also tend to \( \eta_p \) as \( j \to \infty \).

We now approximate those measures with measurable functions. For each \( j = 1, 2, \ldots \), let \( 0 < \varepsilon_j > j^{-2} \) be small enough that the balls \( B_{\varepsilon_j}(x_i^j) \) are disjoint. For \( q \in B_{\varepsilon_j}(x_i^j) \cap \text{supp } \bar{\mu} \) for some \( i \in I_j \), let

\[
g_j(q) = \frac{c_{ij}}{\bar{\mu}(B_{\varepsilon_j}(x_i^j))},
\]

and let \( g_j(q) = 0 \) for all other \( q \in P \). Then \( g_j \bar{\mu} \to \eta_p \), and this proves the lemma. \( \square \)

Proof of Theorem. Assume first that the family \( \mu_s \) exists. To prove that Condition D must hold, let \( f \in C_c^\infty(P) \) be a nonnegative function \( f \in \mathcal{F}_\mu \), and consider the function

\[
g(s) = \int f \, d\mu_s.
\]

Since \( f \) is nonnegative and \( \mu_s \) is a positive measure for all \( s \), \( g \) must be nonnegative as well. Since \( g(0) = 0 \), it must also be true that \( g'(0) = 0 \), and this is equivalent to Condition D.

Now assume that we have a measure \( \mu \) and a distribution \( \eta \) such that Condition D holds, and let us construct a family \( \mu_s \) as in the statement of the theorem. Write \( \eta = \sum_I \partial^I \nu_I \) as in Lemma. The measures \( \nu_I \) induce a measure \( \gamma \) on \( P \) and a family of distributions \( \eta_p \) supported at \( p \in P \) such that for all \( f \in C_c^\infty(P) \)

\[
\eta = \int \eta_p d\gamma(p).
\]

For \( \gamma \)-almost all \( p \), the distributions \( \eta_p \) also satisfy Condition D. From Lemma, we get families \( \mu_p^s \) of measures whose derivatives at 0 are precisely the distributions \( \eta_p \). Thus,

\[
\mu_s = \int \mu_p^s d\gamma(p)
\]

is a family as in the statement of the theorem.

If \( \mu \) is a probability, since each \( \mu_p^s \) preserves the probability, so does \( \mu_s \). \( \square \)
References

[1] 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, second edition, 2008.

[2] Gonzalo Contreras and Renato Iturriaga. *Global minimizers of autonomous Lagrangians*. 22º Colóquio Brasileiro de Matemática. [22nd Brazilian Mathematics Colloquium]. Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 1999.

[3] F. G. Friedlander. *Introduction to the theory of distributions*. Cambridge University Press, Cambridge, second edition, 1998. With additional material by M. Joshi.

[4] John N. Mather. Action minimizing invariant measures for positive definite Lagrangian systems. *Math. Z.*, 207(2):169–207, 1991.

[5] Rodolfo Ríos-Zertuche. The variational structure of the space of holonomic measures. Preprint. arXiv:1408.5785 [math.OC].

[6] O. G. Smolyanov and H. von Weizsäcker. Differentiable families of measures. *J. Funct. Anal.*, 118(2):454–476, 1993.