Uniqueness of the Fisher–Rao metric on the space of smooth densities

Peter W. Michor

University of Vienna, Austria
www.mat.univie.ac.at/~michor

IGAIA IV
Information Geometry and its Applications IV
June 12-17, 2016, Liblice, Czech Republic
In honor of Shun-ichi Amari

Based on:

[M.Bauer, M.Bruveris, P.Michor: Uniqueness of the Fisher–Rao metric on the space of smooth densities, Bull. London Math. Soc. doi:10.1112/blms/bdw020]

[M.Bruveris, P.Michor: Geometry of the Fisher-Rao metric on the space of smooth densities]

[M.Bruveris, P. Michor, A.Parusinski, A. Rainer: Moser's Theorem for manifolds with corners, arxiv:1604.07787]

[M.Bruveris,P.Michor, A.Rainer: Determination of all diffeomorphism invariant tensor fields on the space of smooth positive densities on a compact manifold with corners]

The infinite dimensional geometry used here is based on:

[Andreas Kriegl, Peter W. Michor: The Convenient Setting of Global Analysis. Mathematical Surveys and Monographs, Volume: 53, Amer. Math. Soc., 1997]

Wikipedia [https://en.wikipedia.org/wiki/Convenient_vector_space]

Abstract

For a smooth compact manifold M, any weak Riemannian metric on the space of smooth positive densities which is invariant under the right action of the diffeomorphism group $\mathit{Diff}(M)$ is of the form

$$G_\mu(\alpha, \beta) = C_1(\mu(M)) \int_M \frac{\alpha}{\mu}\frac{\beta}{\mu}\mu + C_2(\mu(M)) \int_M \alpha \cdot \int_M \beta$$

for smooth functions C_1, C_2 of the total volume $\mu(M) = \int_M \mu$.

In this talk the result is extended to:
(0) Geometry of the Fisher-Rao metric: geodesics and curvature.
(1) manifolds with boundary, for manifolds with corner.
(2) to tensor fields of the form $G_\mu(\alpha_1, \alpha_2, \ldots, \alpha_k)$ for any k which are invariant under $\mathit{Diff}(M)$.

The Fisher–Rao metric on the space Prob(M) of probability densities is of importance in the field of information geometry. Restricted to finite-dimensional submanifolds of Prob(M), so-called statistical manifolds, it is called Fisher's information metric [Amari: Differential-geometrical methods in statistics, 1985]. The Fisher–Rao metric is invariant under the action of the diffeomorphism group. A uniqueness result was established [Čencov: Statistical decision rules and optimal inference, 1982, p. 156] for Fisher's information metric on finite sample spaces and [Ay, Jost, Le, Schwachhöfer, 2014] extended it to infinite sample spaces.

The Fisher–Rao metric on the infinite-dimensional manifold of all positive probability densities was studied in [Friedrich: Die Fisher-Information und symplektische Strukturen, 1991], including the computation of its curvature.

The space of densities

Let M^m be a smooth manifold. Let (U_α, u_α) be a smooth atlas for it. The *volume bundle* $(\mathrm{Vol}(M), \pi_M, M)$ of M is the 1-dimensional vector bundle (line bundle) which is given by the following cocycle of transition functions:

$$\psi_{\alpha\beta} : U_{\alpha\beta} = U_\alpha \cap U_\beta \to \mathbb{R} \setminus \{0\} = GL(1, \mathbb{R}),$$

$$\psi_{\alpha\beta}(x) = |\det d(u_\beta \circ u_\alpha^{-1})(u_\alpha(x))| = \frac{1}{|\det d(u_\alpha \circ u_\beta^{-1})(u_\beta(x))|}.$$

Vol(M) is a trivial line bundle over M. But there is no natural trivialization. There is a natural order on each fiber. Since $\mathrm{Vol}(M)$ is a natural bundle of order 1 on M, there is a natural action of the group $\mathrm{Diff}(M)$ on $\mathrm{Vol}(M)$, given by

$$\begin{array}{ccc} \mathrm{Vol(M)} & \xrightarrow{|\det(T\varphi^{-1})| \circ \varphi} & \mathrm{Vol(M)} \\ \big\downarrow & & \big\downarrow \\ M & \xrightarrow{\varphi} & M \end{array}.$$

If M is orientable, then $\mathrm{Vol}(M) = \Lambda^m T^*M$. If M is not orientable, let $\tilde{M}$ be the orientable double cover of M with its deck-transformation $\tau : \tilde{M} \to \tilde{M}$. Then $\Gamma(\mathrm{Vol}(M))$ is isomorphic to the space $\{\omega \in \Omega^m(\tilde{M}) : \tau^*\omega = -\omega\}$. These are the 'formes impaires' of de Rham. See [M 2008, 13.1] for this.

Sections of the line bundle $\mathrm{Vol}(M)$ are called densities. The space $\Gamma(\mathrm{Vol}(M))$ of all smooth sections is a Fréchet space in its natural topology; see [Kriegl-M, 1997]. For each section α of $\mathrm{Vol}(M)$ of compact support the integral $\int_M \alpha$ is invariantly defined as follows: Let (U_α, u_α) be an atlas on M with associated trivialization $\psi_\alpha : \mathrm{Vol}(M)|_{U_\alpha} \to \mathbb{R}$, and let f_α be a partition of unity with $\mathrm{supp}(f_\alpha) \subset U_\alpha$. Then we put

$$\int_M \mu = \sum_\alpha \int_{U_\alpha} f_\alpha \mu := \sum_\alpha \int_{u_\alpha(U_\alpha)} f_\alpha(u_\alpha^{-1}(y)).\psi_\alpha(\mu(u_\alpha^{-1}(y)))\, dy.$$

The integral is independent of the choice of the atlas and the partition of unity.

The Fisher–Rao metric

Let M^m be a smooth compact manifold without boundary. Let $\mathrm{Dens}_+(M)$ be the space of smooth positive densities on M, i.e., $\mathrm{Dens}_+(M) = \{\mu \in \Gamma(\mathrm{Vol}(M)) \,:\, \mu(x) > 0 \; \forall x \in M\}$.
Let $\mathrm{Prob}(M)$ be the subspace of positive densities with integral 1.
For $\mu \in \mathrm{Dens}_+(M)$ we have $T_\mu \,\mathrm{Dens}_+(M) = \Gamma(\mathrm{Vol}(M))$ and for $\mu \in \mathrm{Prob}(M)$ we have
$T_\mu \,\mathrm{Prob}(M) = \{\alpha \in \Gamma(\mathrm{Vol}(M)) : \int_M \alpha = 0\}$.
The Fisher–Rao metric on $\mathrm{Prob}(M)$ is defined as:

$$G^{\mathrm{FR}}_\mu(\alpha, \beta) = \int_M \frac{\alpha}{\mu}\frac{\beta}{\mu}\mu.$$

It is invariant for the action of $\mathrm{Diff}(M)$ on $\mathrm{Prob}(M)$:

$$\Big((\varphi^*)^* G^{\mathrm{FR}}\Big)_\mu(\alpha, \beta) = G^{\mathrm{FR}}_{\varphi^*\mu}(\varphi^*\alpha, \varphi^*\beta) =$$

$$= \int_M \Big(\frac{\alpha}{\mu} \circ \varphi\Big)\Big(\frac{\beta}{\mu} \circ \varphi\Big)\varphi^*\mu = \int_M \frac{\alpha}{\mu}\frac{\beta}{\mu}\mu\,.$$

Theorem [BBM, 2016]

Let M be a compact manifold without boundary of dimension ≥ 2. Let G be a smooth (equivalently, bounded) bilinear form on $\mathrm{Dens}_+(M)$ *which is invariant under the action of* $\mathrm{Diff}(M)$*. Then*

$$G_\mu(\alpha,\beta) = C_1(\mu(M)) \int_M \frac{\alpha}{\mu}\frac{\beta}{\mu}\,\mu + C_2(\mu(M)) \int_M \alpha \cdot \int_M \beta$$

for smooth functions C_1, C_2 of the total volume $\mu(M)$.

To see that this theorem implies the uniqueness of the Fisher–Rao metric, note that if G is a $\mathrm{Diff}(M)$-invariant Riemannian metric on $\mathrm{Prob}(M)$, then we can equivariantly extend it to $\mathrm{Dens}_+(M)$ via

$$G_\mu(\alpha,\beta) = G_{\frac{\mu}{\mu(M)}}\left(\alpha - \Big(\int_M \alpha\Big)\frac{\mu}{\mu(M)}, \beta - \Big(\int_M \beta\Big)\frac{\mu}{\mu(M)}\right) .$$

Relations to right-invariant metrics on diffeom. groups

Let $\mu_0 \in \operatorname{Prob}(M)$ be a fixed smooth probability density. In [Khesin, Lenells, Misiolek, Preston, 2013] it has been shown, that the degenerate, $\dot{H}^1$-metric $\frac{1}{2}\int_M \operatorname{div}^{\mu_0}(X). \operatorname{div}^{\mu_0}(X).\mu_0$ on $\mathfrak{X}(M)$ is invariant under the adjoint action of $\operatorname{Diff}(M, \mu_0)$. Thus the induced degenerate right invariant metric on $\operatorname{Diff}(M)$ descends to a metric on $\operatorname{Prob}(M) \cong \operatorname{Diff}(M, \mu_0)\backslash \operatorname{Diff}(M)$ via

$$\operatorname{Diff}(M) \ni \varphi \mapsto \varphi^* \mu_0 \in \operatorname{Prob}(M)$$

which is invariant under the right action of $\operatorname{Diff}(M)$. This is the Fisher–Rao metric on $\operatorname{Prob}(M)$. In [Modin, 2014], the $\dot{H}^1$-metric was extended to a non-degenerate metric on $\operatorname{Diff}(M)$, also descending to the Fisher–Rao metric.

Corollary. *Let* $\dim(M) \geq 2$*. If a weak right-invariant (possibly degenerate) Riemannian metric* $\tilde{G}$ *on* $\mathrm{Diff}(M)$ *descends to a metric* G *on* $\mathrm{Prob}(M)$ *via the right action, i.e., the mapping* $\varphi \mapsto \varphi^* \mu_0$ *from* $(\mathrm{Diff}(M), \tilde{G})$ *to* $(\mathrm{Prob}(M), G)$ *is a Riemannian submersion, then* G *has to be a multiple of the Fisher–Rao metric.*

Note that any right invariant metric $\tilde{G}$ on $\mathrm{Diff}(M)$ descends to a metric on $\mathrm{Prob}(M)$ via $\varphi \mapsto \varphi_* \mu_0$; but this is not $\mathrm{Diff}(M)$-invariant in general.

Invariant metrics on $\mathrm{Dens}_+(S^1)$.

$\mathrm{Dens}_+(S^1) = \Omega^1_+(S^1)$, and $\mathrm{Dens}_+(S^1)$ is $\mathrm{Diff}(S^1)$-equivariantly isomorphic to the space of all Riemannian metrics on S^1 via $\Phi = (\quad)^2 : \mathrm{Dens}_+(S^1) \to \mathrm{Met}(S^1)$, $\Phi(fd\theta) = f^2 d\theta^2$.
On $\mathrm{Met}(S^1)$ there are many $\mathrm{Diff}(S^1)$-invariant metrics; see [Bauer, Harms, M, 2013]. For example Sobolev-type metrics. Write $g \in \mathrm{Met}(S^1)$ in the form $g = \tilde{g} d\theta^2$ and $h = \tilde{h} d\theta^2$, $k = \tilde{k} d\theta^2$ with $\tilde{g}, \tilde{h}, \tilde{k} \in C^\infty(S^1)$. The following metrics are $\mathrm{Diff}(S^1)$-invariant:

$$G^l_g(h,k) = \int_{S^1} \frac{\tilde{h}}{\tilde{g}}.\,(1+\Delta^g)^n \left(\frac{\tilde{k}}{\tilde{g}}\right) \sqrt{\tilde{g}}\, d\theta\,;$$

here Δ^g is the Laplacian on S^1 with respect to the metric g. The pullback by Φ yields a $\mathrm{Diff}(S^1)$-invariant metric on $\mathrm{Dens}_+(M)$:

$$G_\mu(\alpha,\beta) = 4\int_{S^1} \frac{\alpha}{\mu}.\left(1+\Delta^{\Phi(\mu)}\right)^n \left(\frac{\beta}{\mu}\right) \mu\,.$$

For $n=0$ this is 4 times the Fisher–Rao metric. For $n \geq 1$ we get different $\mathrm{Diff}(S^1)$-invariant metrics on $\mathrm{Dens}_+(M)$ and on $\mathrm{Prob}(S^1)$.

Main Theorem

Let M be a compact manifold, possibly with corners, of dimension ≥ 2. Let G be a smooth (equivalently, bounded) $\binom{0}{n}$-tensor field on $\mathrm{Dens}_+(M)$ which is invariant under the action of $\mathrm{Diff}(M)$. If M is not orientable or if $n \leq \dim(M) = m$, then

$$
\begin{aligned}
&G_\mu(\alpha_1, \dots, \alpha_n) = C_0(\mu(M)) \int_M \frac{\alpha_1}{\mu} \dots \frac{\alpha_n}{\mu}\, \mu \\
&+ \sum_{i=1}^{n} C_i(\mu(M)) \int_M \alpha_i \cdot \int_M \frac{\alpha_1}{\mu} \dots \frac{\widehat{\alpha_i}}{\mu} \dots \frac{\alpha_n}{\mu}\, \mu \\
&+ \sum_{i<j}^{n} C_{ij}(\mu(M)) \int_M \frac{\alpha_i}{\mu} \frac{\alpha_j}{\mu}\, \mu \cdot \int_M \frac{\alpha_1}{\mu} \dots \frac{\widehat{\alpha_i}}{\mu} \dots \frac{\widehat{\alpha_i}}{\mu} \dots \frac{\alpha_n}{\mu}\, \mu \\
&+ \dots \\
&+ C_{12\dots n}(\mu(M)) \int_M \frac{\alpha_1}{\mu}\, \mu \cdot \int_M \frac{\alpha_2}{\mu}\, \mu \cdot \dots \cdot \int_M \frac{\alpha_n}{\mu}\, \mu.
\end{aligned}
$$

for some smooth functions $C_0, \dots$ of the total volume $\mu(M)$.

Main Theorem, continued

If M is orientable and $n > \dim(M) = m$, then each integral over more than m functions α_i/μ has to be replaced by the following expression which we write only for the first term:

$$C_0(\mu(M)) \int_M \frac{\alpha_1}{\mu} \dots \frac{\alpha_n}{\mu}\, \mu +$$

$$+ \sum C_0^K(\mu(M)) \int \frac{\alpha_{k_1}}{\mu} \dots \frac{\alpha_{k_{n-m}}}{\mu}\, d\big(\frac{\alpha_{k_{n-m+1}}}{\mu}\big) \wedge \dots \wedge d\big(\frac{\alpha_{k_n}}{\mu}\big)$$

where $K = \{k_{n-m+1}, \dots, k_n\}$ runs through all subsets of $\{1, \dots, n\}$ containing exactly m elements.

Moser's theorem for manifolds with corners [BMPR16]

Let M be a compact smooth manifold with corners, possibly non-orientable. Let μ_0 and μ_1 be two smooth positive densities in $\mathrm{Dens}_+(M)$ *with* $\int_M \mu_0 = \int_M \mu_1$. *Then there exists a diffeomorphism* $\varphi : M \to M$ *such that* $\mu_1 = \varphi^* \mu_0$. *If and only if* $\mu_0(x) = \mu_1(x)$ *for each corner* $x \in \partial^{\geq 2} M$ *of codimension* ≥ 2, *then* φ *can be chosen to be the identity on* ∂M.

This result is highly desirable even for M a simplex. The proof is essentially contained in [Banyaga1974], who proved it for manifolds with boundary.

Geometry of the Fisher-Rao metric

$$G_\mu(\alpha,\beta) = C_1(\mu(M)) \int_M \frac{\alpha}{\mu}\frac{\beta}{\mu}\,\mu + C_2(\mu(M)) \int_M \alpha \cdot \int_M \beta$$

This metric will be studied in different representations.

$$\mathrm{Dens}_+(M) \xrightarrow{R} C^\infty(M,\mathbb{R}_{>0}) \xrightarrow{\Phi} \mathbb{R}_{>0} \times S \cap C^\infty_{>0} \xrightarrow{W\times \mathrm{Id}} (W_-,W_+) \times S \cap C^\infty_{>0}\,.$$

We fix $\mu_0 \in \mathrm{Prob}(M)$ and consider the mapping

$$R : \mathrm{Dens}_+(M) \to C^\infty(M,\mathbb{R}_{>0})\,, \qquad R(\mu) = f = \sqrt{\frac{\mu}{\mu_0}}\,.$$

The map R is a diffeomorphism and we will denote the induced metric by $\tilde{G} = \left(R^{-1}\right)^* G$; it is given by the formula

$$\tilde{G}_f(h,k) = 4C_1(\|f\|^2)\langle h,k\rangle + 4C_2(\|f\|^2)\langle f,h\rangle\langle f,k\rangle\,,$$

and this formula makes sense for $f \in C^\infty(M,\mathbb{R}) \setminus \{0\}$.
The map R is inspired by [B. Khesin, J. Lenells, G. Misiolek, S. C. Preston: Geometry of diffeomorphism groups, complete integrability and geometric statistics. Geom. Funct. Anal., 23(1):334-366, 2013.]

Remark on R^{-1}

$$R^{-1} : C^\infty(M, \mathbb{R}) \to \Gamma_{\geq 0}(\mathrm{Vol}(M)), \quad f \mapsto f^2 \mu_0$$

makes sense on the whole space $C^\infty(M, \mathbb{R})$ and its image is stratified (loosely speaking) according to the rank of TR^{-1}. The image looks somewhat like the orbit space of a discrete reflection group. Geodesics are mapped to curves which are geodesics in the interior $\Gamma_{>0}(\mathrm{Vol}(M))$, and they are reflected following Snell's law at some hyperplanes in the boundary.

Polar coordinates

on the pre-Hilbert space $(C^\infty(M,\mathbb{R}), \langle\ ,\ \rangle_{L^2(\mu_0)})$. Let $S = \{\varphi \in L^2(M,\mathbb{R}) : \int_M \varphi^2 \mu_0 = 1\}$ denote the L^2-sphere. Then

$$\Phi : C^\infty(M,\mathbb{R})\backslash\{0\} \to \mathbb{R}_{>0}\times(S\cap C^\infty)\,, \qquad \Phi(f) = (r,\varphi) = \left(\|f\|, \frac{f}{\|f\|}\right)$$

is a diffeomorphism. We set $\bar{G} = \left(\Phi^{-1}\right)^* \tilde{G}$; the metric has the expression

$$\bar{G}_{r,\varphi} = g_1(r)\langle d\varphi, d\varphi\rangle + g_2(r)dr^2\,,$$

with $g_1(r) = 4C_1(r^2)r^2$ and $g_2(r) = 4\left(C_1(r^2) + C_2(r^2)r^2\right)$. Finally we change the coordinate r diffeomorphically to

$$s = W(r) = 2\int_1^r \sqrt{g_2(\rho)}\,d\rho\,.$$

Then, defining $a(s) = 4C_1(r(s)^2)r(s)^2$, we have

$$\boxed{\bar{G}_{s,\varphi} = a(s)\langle d\varphi, d\varphi\rangle + ds^2\,.}$$

Let $W_- = \lim_{r\to 0+} W(r)$ and $W_+ = \lim_{r\to\infty} W(r)$. Then $W : \mathbb{R}_{>0} \to (W_-, W_+)$ is a diffeomorphism.

This completes the first row in Fig. 1.

$$\begin{array}{ccccccc}
\mathrm{Dens}_+(M) & \xrightarrow{R} & C^\infty(M,\mathbb{R}_{>0}) & \xrightarrow{\Phi} & \mathbb{R}_{>0}\times S\cap C^\infty_{>0} & \xrightarrow{W\times\mathrm{Id}} & (W_-,W_+)\times S\cap C^\infty_{>0} \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
\mathrm{Dens}(M)\setminus\{0\} & \xrightarrow{R} & C^0(M,\mathbb{R})\setminus\{0\} & \xrightarrow{\Phi} & \mathbb{R}_{>0}\times S\cap C^0 & \xrightarrow{W\times\mathrm{Id}} & \mathbb{R}\times S\cap C^0 \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
\Gamma_{L^1}(\mathrm{Vol}(M))\setminus\{0\} & \xrightarrow{R} & L^2(M,\mathbb{R})\setminus\{0\} & \xrightarrow{\Phi} & \mathbb{R}_{>0}\times S & \xrightarrow{W\times\mathrm{Id}} & \mathbb{R}\times S
\end{array}$$

Figure: Representations of $\mathrm{Dens}_+(M)$ and its completions. In the second and third rows we assume that $(W_-, W_+) = (-\infty, +\infty)$ and we note that R is a diffeomorphism only in the first row.

Geodesic equation:

$$\boxed{\begin{aligned}
\nabla^S_{\partial_t}\varphi_t &= \partial_t\left(\log g_1(r)\right)\varphi_t \\
r_{tt} &= \frac{C_0^2}{2}\frac{g_1'(r)}{g_1(r)^2 g_2(r)} - \frac{1}{2}\partial_t\left(\log g_2(r)\right) r_t
\end{aligned}}$$

Since $\bar{G}$ induces the canonical metric on (W_-, W_+), a necessary condition for $\bar{G}$ to be complete is $(W_-, W_+) = (-\infty, +\infty)$. Rewritten in terms of the functions C_1, C_2 this becomes

$$W_+ = \infty \Leftrightarrow \left(\int_1^\infty r^{-1/2} \sqrt{C_1(r)}\, dr = \infty \text{ or } \int_1^\infty \sqrt{C_2(r)}\, dr = \infty \right),$$

and similarly for $W_- = -\infty$, with the limits of the integration being 0 and 1.

Relation to hypersurfaces of revolution in the (pre-) Hilbert space

We consider the metric on $(W_-, W_+) \times S \cap C^\infty$ in the form $\tilde{G}_{r,\varphi} = a(s)\langle d\varphi, d\varphi\rangle + ds^2$ where $a(s) = 4C_1(r(s)^2)r(s)^2$. Then we consider the isometric embedding (remember $\langle \varphi, d\varphi\rangle = 0$ on $S \cap C^\infty$)

$$\Psi : ((W_-, W_+) \times S \cap C^\infty, \tilde{G}) \to \left(\mathbb{R} \times C^\infty(M, \mathbb{R}), du^2 + \langle df, df\rangle\right),$$

$$\Psi(s, \varphi) = \left(\int_0^s \sqrt{1 - \frac{a'(\sigma)^2}{4a(\sigma)}}\, d\sigma \;,\; \sqrt{a(s)}\varphi \right),$$

which defined and smooth only on the open subset

$$R := \{(s, \varphi) \in (W_-, W_+) \times S \cap C^\infty : a'(s)^2 < 4a(s)\}.$$

Fix some $\varphi_0 \in S \cap C^\infty$ and consider the generating curve

$$s \mapsto \left(\int_0^s \sqrt{1 - \frac{a'(\sigma)^2}{4a(\sigma)}}\, d\sigma \;,\; \sqrt{a(s)} \right) \in \mathbb{R}^2 .$$

Then s is an arc-length parameterization of this curve!

Given any arc-length parameterized curve $I \ni s \mapsto (c_1(s), c_2(s))$ in $\mathbb{R}^2$ and its generated hypersurface of rotation

$$\{(c_1(s), c_2(s)\varphi) : s \in I, \varphi \in S \cap C^\infty\} \subset \mathbb{R} \times C^\infty(M, \mathbb{R}),$$

the induced metric in the (s, φ)-parameterization is $ds^2 + c_2(s)^2 \langle d\varphi, d\varphi \rangle$.

This suggests that the moduli space of hypersurfaces of revolution is naturally embedded in the moduli space of all metrics of the form (b).

Theorem

If $(W_-, W_+) = (-\infty, +\infty)$, *then any two points* (s_0, φ_0) *and* (s_1, φ_1) *in* $\mathbb{R} \times S$ *can be joined by a minimal geodesic. If* φ_0 *and* φ_1 *lie in* $S \cap C^\infty$, *then the minimal geodesic lies in* $\mathbb{R} \times S \cap C^\infty$.

Proof. If φ_0 and φ_1 are linearly independent, we consider the 2-space $V = V(\varphi_0, \varphi_1)$ spanned by φ_0 and φ_1 in L^2. Then $\mathbb{R} \times V \cap S$ is totally geodesic since it is the fixed point set of the isometry $(s, \varphi) \mapsto (s, \mathfrak{s}_V(\varphi))$ where $\mathfrak{s}_V$ is the orthogonal reflection at V. Thus there is exists a minimizing geodesic between (s_0, φ_0) and (s_1, φ_1) in the complete 3-dimensional Riemannian submanifold $\mathbb{R} \times V \cap S$. This geodesic is also length-minimizing in the strong Hilbert manifold $\mathbb{R} \times S$ by the following arguments:

Given any smooth curve $c = (s, \varphi) : [0, 1] \to \mathbb{R} \times S$ between these two points, there is a subdivision $0 = t_0 < t_1 < \cdots < t_N = 1$ such that the piecewise geodesic c_1 which first runs along a geodesic from $c(t_0)$ to $c(t_1)$, then to $c(t_2)$, ..., and finally to $c(t_N)$, has length $\mathrm{Len}(c_1) \le \mathrm{Len}(c)$. This piecewise geodesic now lies in the totally geodesic $(N+2)$-dimensional submanifold $\mathbb{R} \times V(\varphi(t_0), \dots, \varphi(t_N)) \cap S$. Thus there exists a geodesic c_2 between the two points (s_0, φ_0) and s_1, φ_1 which is length minimizing in this $(N+2)$-dimensional submanifold. Therefore $\mathrm{Len}(c_2) \le \mathrm{Len}(c_1) \le \mathrm{Len}(c)$. Moreover, $c_2 = (s \circ c_2, \varphi \circ c_2)$ lies in $\mathbb{R} \times V(\varphi_0, (\varphi \circ c_2)'(0)) \cap S$ which also contains φ_1, thus c_2 lies in $\mathbb{R} \times V(\varphi_0, \varphi_1) \cap S$.

If $\varphi_0 = \varphi_1$, then $\mathbb{R} \times \{\varphi_0\}$ is a minimal geodesic. If $\varphi_0 = -\varphi_0$ we choose a great circle between them which lies in a 2-space V and proceed as above. □

Covariant derivative

On $\mathbb{R} \times S$ (we assume that $(W_-, W_+) = \mathbb{R}$) with metric $\bar{G} = ds^2 + a(s)\langle d\varphi, d\varphi\rangle$ we consider smooth vector fields $f(s,\varphi)\partial_s + X(s,\varphi)$ where $X(s,\ ) \in \mathfrak{X}(S)$ is a smooth vector field on the Hilbert sphere S. We denote by ∇^S the covariant derivative on S and get

$$\boxed{\begin{aligned}\nabla_{f\partial_s + X}(g\partial_s + Y) &= \Big(f.g_s + dg(X) - \frac{a_s}{2}\langle X, Y\rangle\Big)\partial_s \\ &\quad + \frac{a_s}{2a}(fY + gX) + fY_s + \nabla^S_X Y\end{aligned}}$$

Curvature:

$$\boxed{\begin{aligned}&\mathcal{R}(f\partial_s + X, g\partial_s + Y)(h\partial_s + Z) = \\ &= \Big(\frac{a_{ss}}{2} - \frac{a_s^2}{4a}\Big)\langle gX - fY, Z\rangle\partial_s + \mathcal{R}^S(X,Y)Z \\ &\quad - \Big(\Big(\frac{a_s}{2a}\Big)_s + \frac{a_s^2}{4a^2}\Big)h(gX - fY) + \frac{a_s}{2a}\big(\langle X, Z\rangle Y - \langle Y, Z\rangle X\big)\,.\end{aligned}}$$

Sectional Curvature

Let us take $X, Y \in T_\varphi S$ with $\langle X, Y\rangle = 0$ and $\langle X, X\rangle = \langle Y, Y\rangle = 1/a(s)$, then

$$\boxed{\begin{aligned} \mathrm{Sec}_{(s,\varphi)}(\mathrm{span}(X, Y)) &= \frac{1}{a} - \frac{a_s}{2a^2}, \\ \mathrm{Sec}_{(s,\varphi)}(\mathrm{span}(\partial_s, Y)) &= -\frac{a_{ss}}{2a} + \frac{a_s^2}{4a^2} \end{aligned}}$$

are all the possible sectional curvatures.

Back to the Main Theorem

Let M be a compact manifold, possibly with corners, of dimension ≥ 2. Then the space of all $\mathrm{Diff}(M)$-invariant purely covariant tensor fields on $\mathrm{Dens}_+(M)$ is generated as algebra with unit 1 over the ring of of smooth functions $f(\mu(M))$, $f \in C^\infty(\mathbb{R}, \mathbb{R})$ by the following generators, allowing for permutations of the entries $\alpha_i \in T_\mu \mathrm{Dens}_+(M)$:

$$\int_M \frac{\alpha_1}{\mu} \dots \frac{\alpha_n}{\mu}\, \mu \qquad \textit{for all } n \in \mathbb{N}_{>0}, \textit{ and by}$$

$$\int \frac{\alpha_1}{\mu} \dots \frac{\alpha_{n-m}}{\mu}\, d\Big(\frac{\alpha_{n-m+1}}{\mu}\Big) \wedge \dots \wedge d\Big(\frac{\alpha_n}{\mu}\Big)$$

for $n > \dim(M)$ and orientable M.

Manifolds with corners alias quadrantic (orthantic) manifolds

For more information we refer to [DouadyHerault73], [Michor80], [Melrose96], etc. Let $Q = Q^m = \mathbb{R}^m_{\geq 0}$ be the positive orthant or quadrant. By Whitney's extension theorem or Seeley's theorem, restriction $C^\infty(\mathbb{R}^m) \to C^\infty(Q)$ is a surjective continuous linear mapping which admits a continuous linear section (extension mapping); so $C^\infty(Q)$ is a direct summand in $C^\infty(\mathbb{R}^m)$. A point $x \in Q$ is called a *corner of codimension* $q > 0$ if x lies in the intersection of q distinct coordinate hyperplanes. Let $\partial^q Q$ denote the set of all corners of codimension q.

A manifold with corners (recently also called a quadrantic manifold) M is a smooth manifold modelled on open subsets of Q^m. We assume that it is connected and second countable; then it is paracompact and for each open cover it admits a subordinated smooth partition of unity. Any manifold with corners M is a submanifold with corners of an open manifold $\tilde{M}$ of the same dim. Restriction $C^\infty(\tilde{M}) \to C^\infty(M)$ is a surjective continuous linear map which admits a continuous linear section.Thus $C^\infty(M)$ is a topological direct summand in $C^\infty(\tilde{M})$ and the same holds for the dual spaces: The space of distributions $\mathcal{D}'(M)$, which we identity with $C^\infty(M)'$, is a direct summand in $\mathcal{D}'(\tilde{M})$. It consists of all distributions with support in M.

We do not assume that M is oriented, but eventually we will assume that M is compact. Diffeomorphisms of M map the boundary ∂M to itself and map the boundary $\partial^q M$ of corners of codimension q to itself; $\partial^q M$ is a submanifold of codimension q in M; in general $\partial^q M$ has finitely many connected components. We shall consider ∂M as stratified into the connected components of all $\partial^q M$ for $q > 0$.

Beginning of the proof of the Main Theorem

Fix a basic probability density μ_0. By Moser's theorem for manifolds with corners, for each $\mu \in \text{Dens}_+(M)$ there exists a diffeomorphism $\varphi_\mu \in \text{Diff}(M)$ with $\varphi_\mu^* \mu = \mu(M)\mu_0 =: c.\mu_0$ where $c = \mu(M) = \int_M \mu > 0$. Then

$$\big((\varphi_\mu^*)^* G\big)_\mu(\alpha_1, \dots, \alpha_n) = G_{\varphi_\mu^* \mu}(\varphi_\mu^* \alpha_1, \dots, \varphi_\mu^* \alpha_n) =$$
$$= G_{c.\mu_0}(\varphi_\mu^* \alpha_1, \dots, \varphi_\mu^* \alpha_n)\,.$$

Thus it suffices to show that for any $c > 0$ we have

$$G_{c\mu_0}(\alpha_1, \dots, \alpha_n) = C_0(c).\int_M \frac{\alpha_1}{\mu_0} \dots \frac{\alpha_n}{\mu_0} \mu_0 + \dots$$

for some functions $C_0, \dots$ of the total volume $c = \mu(M)$. Since $c \mapsto c.\mu_0$ is a smooth curve in $\text{Dens}_+(M)$, the functions $C_0, \dots$ are then smooth in c. All k-linear forms are still invariant under the action of the group
$\text{Diff}(M, c\mu_0) = \text{Diff}(M, \mu_0) = \{\psi \in \text{Diff}(M) : \psi^* \mu_0 = \mu_0\}$.

The k-linear form

$$\big(T_{\mu_0}\operatorname{Dens}_+(M)\big)^k \ni (\alpha_1,\dots,\alpha_n) \mapsto G_{c\mu_0}\Big(\frac{\alpha_1}{\mu_0}\mu_0,\dots,\frac{\alpha_n}{\mu_0}\mu_0\Big)$$

can be viewed as a bounded k-linear form

$$C^\infty(M)^k \ni (f_1,\dots,f_n) \mapsto G_c(f_1,\dots,f_n)\,.$$

Using the Schwartz kernel theorem, $\check{G}_c$ has a kernel $\hat{G}_c$, which is a distribution (generalized function) in

$$\begin{aligned}\mathcal{D}'(M^n) \cong \mathcal{D}'(M)\,\bar{\otimes}\dots\bar{\otimes}\,\mathcal{D}'(M) &= \big(C^\infty(M)\,\bar{\otimes}\dots\bar{\otimes}\,C^\infty(M)\big)' \\ &\cong L(C^\infty(M^k),\mathcal{D}'(M^{n-k}))\,.\end{aligned}$$

Note the defining relations

$$G_c(f_1,\dots,f_n) = \langle \check{G}_c(f_1,\dots,f_k), f_{k+1}\otimes\dots\otimes f_n\rangle = \langle \hat{G}_c, f_1\otimes\dots\otimes f_n\rangle\,.$$

$\hat{G}_c$ is invariant under the diagonal action of $\operatorname{Diff}(M,\mu_0)$ on M^n.

The infinitesimal version of this invariance is:

$$
\begin{aligned}
0 &= \langle \mathcal{L}_{X^{\text{diag}}} \hat{G}_c, f_1 \otimes \cdots \otimes f_n \rangle = -\langle \hat{G}_c, \mathcal{L}_{X^{\text{diag}}} (f_1 \otimes \cdots \otimes f_n) \rangle \\
&= -\sum_{i=1}^{n} \langle \hat{G}_c, f_1 \otimes \cdots \otimes \mathcal{L}_X f_i \otimes \cdots \otimes f_n) \rangle \\
X^{\text{diag}} &= X \times 0 \times \ldots \times 0 + 0 \times X \times 0 \times \ldots \times 0 + \ldots .
\end{aligned}
$$

for all $X \in \mathfrak{X}(M, \mu_0)$.

We will consider various (permuted versions) of the associated bounded mappings

$$
\check{G}_c : C^\infty(M)^k \to \big(C^\infty(M)^{n-k} \big)' = \mathcal{D}'(M^{n-k}) \, .
$$

We shall use the fixed density $\mu_0 \in \text{Dens}_+(M)$ for the rest of this section. So we identify distributions on M^k with the dual space $C^\infty(M^k)' =: \mathcal{D}'(M^k)$

The Lie algebra of $\mathrm{Diff}(M, \mu_0)$

For a fixed positive density μ_0 on M, the Lie algebra of $\mathrm{Diff}(M, \mu_0)$ which we will denote by $\mathfrak{X}(M, \partial M, \mu_0)$, is the subalgebra of vector fields which are tangent to each boundary stratum and which are divergence free: $0 = \mathrm{div}^{\mu_0}(X) := \frac{\mathcal{L}_X \mu_0}{\mu_0}$. These are exactly the fields X such that for each good subset U (where each density can be identified with an m-form) the form $\hat{\imath}_{\mu_0}(X)$ is a closed form in $\Omega^{m-1}(U, \partial U)$, and $0 = \mathrm{div}^{\mu_0}(X) := \frac{\mathcal{L}_X \mu_0}{\mu_0}$.

Denote by $\mathfrak{X}_{\mathrm{exact}}(M, \partial M, \mu_0)$ the set (not a vector space) of 'exact' divergence free vector fields $X = \hat{\imath}_{\mu_0}^{-1}(d\omega)$, where $\omega \in \Omega_c^{m-2}(U, \partial U)$ for a good subset $U \subset M$. They are automatically tangent to each boundary stratum since $d\omega \in \Omega_c^{m-1}(U, \partial U)$.

Lemma *If for* $f \in C^\infty(M)$ *and a good set* $U \subseteq M$ *we have* $(\mathcal{L}_X f)|U = 0$ *for all* $X \in \mathfrak{X}_{exact}(M, \partial M, \mu_0)$*, then* $f|U$ *is constant.*

Lemma *If for a distribution* $A \in \mathcal{D}'(M) = C^\infty(M)'$ *and a connected open set* $U \subseteq M$ *we have* $\mathcal{L}_X A|U = 0$ *for all* $X \in \mathfrak{X}_{exact}(M, \partial M, \mu_0)$*, then* $A|U = C\mu_0|U$ *for some constant* C*, meaning* $\langle A, f\rangle = C\int_M f\mu_0$ *for all* $f \in C_c^\infty(U)$*.*

This lemma proves the theorem for the case $n = 1$.

Lemma *Each operator*

$$\check{G}_c : C^\infty(M) \to \mathcal{C}^\infty(M^{n-1})'$$
$$f_i \mapsto \big((f_1, \dots \widehat{f_i} \dots, f_n) \mapsto G_c(f_1, \dots, f_n)\big)$$

has the following property: If for $f \in C^\infty(M)$ *and a connected open* $U \subseteq M$ *the restriction* $f|U$ *is constant, then* $\mathcal{L}_{X^{diag}}(\check{G}_c(f))|U^{n-1} = 0$ *for each exact vector field* $X \in \mathfrak{X}_{exact}(M, \partial M, \mu_0)$*.*

Lemma *Let* $\hat{G}$ *be an invariant distribution in* $\mathcal{D}'(M^n)$*. Then for each* $1 \leq i \leq n$ *there exists an invariant distribution* $\hat{G}_i \in \mathcal{D}'(M^{n-1})$ *such that the distribution*

$$(f_1, \ldots, f_n) \mapsto \hat{G}(f_1, \ldots, f_n) - \hat{G}_i(f_1, \ldots \widehat{f_i} \ldots, f_n) \cdot \int_M f_i \mu_0$$

has support in the set

$$D_i(M) = \{(x_1, \ldots, x_n) : x_i = x_j \text{ for some } j \neq i\} .$$

Lemma *There exists a constant* $C = C(c)$ *such that the distribution* $\hat{G}_c - C{\mu_0}^{\otimes n}$ *is supported on the union of all partial diagonals*

$$D := \{(x_1, \ldots, x_n) \in M^n \ : \ \text{for at least one pair } i \neq j \text{ we have equality: } x_i = x_j\} .$$

Lemma *Let* $\hat{G} \in \mathcal{D}'(M^n)$ *be a* $\mathrm{Diff}(M, \mu_0)$*-invariant distribution, supported on the full diagonal* $\Delta(M) = \{(x_1, \ldots, x_n) \in M^n \, : \, x_1 = \cdots = x_n\} \subset M^n$*. If* $n \leq \dim(M)$ *or if* M *is not orientable, there exist some constant* C *such that* $G(f_1, \ldots, f_n) = C \int_M f_1 \ldots f_n \mu_0$*.*
If $n > \dim(M)$ *and if* M *is orientable, then there exist constants such that*

$$C_0 \int_M \frac{\alpha_1}{\mu} \ldots \frac{\alpha_n}{\mu} \mu + \\ + \sum C_0^K \int \frac{\alpha_{k_1}}{\mu} \ldots \frac{\alpha_{k_{n-m}}}{\mu} d\Big(\frac{\alpha_{k_{n-m+1}}}{\mu}\Big) \wedge \cdots \wedge d\Big(\frac{\alpha_{k_n}}{\mu}\Big)$$

where $K = \{k_{n-m+1}, \ldots, k_n\}$ *runs through all subsets of* $\{1, \ldots, n\}$ *containing exactly* m *elements.*

Beginning of the proof of the lemma:

Let (U, u) be an oriented chart on M, diffeomorphic to Q_p^m with coordinates $u^1 \geq 0, \ldots, u^p \geq 0, u^{p+1}, \ldots, u^m$, such that $\mu_0 | U = du^1 \wedge \cdots \wedge du^m$. The distribution $\hat{G}|_U \in D'(U^n)$ has support contained in the full diagonal $\Delta(U) = \{(x, \ldots, x) \in U^n \,:\, x \in U\}$ and is of finite order k since M is compact. By Thm. 2.3.5 of Hörmander 1983, the corresponding multilinear form G can be written as

$$G(f_1, \ldots, f_n) = \sum_{|\alpha_1| + \ldots + |\alpha_{n-1}| \leq k} \left\langle A_{\alpha_1, \ldots, \alpha_{n-1}}, \partial^{\alpha_1} f_1 \ldots \partial^{\alpha_{n-1}} f_{n-1} . f_n \right\rangle \ ,$$

with multi-indices $\alpha_j = (\alpha_{j,1}, \ldots, \alpha_{j,m})$ and unique distributions $A_{\alpha_1, \ldots, \alpha_{n-1}} \in D'(U)$ of order $k - |\alpha_1| - \ldots - |\alpha_{n-1}|$.

End of the proof of the Main Theorem

Let $\hat{G}$ be an invariant distribution in $\mathcal{D}'(M^n)$ and let $k < n/2$. Let $\{1, \ldots, n\} = \{i_1, \ldots, i_k\} \sqcup \{j_1, \ldots, j_{n-k}\}$ be a partition into a disjoint union.

Without loss, let $\{i_1, \ldots, i_k\} = \{1, \ldots, k\}$. Let $(x_1, \ldots, x_n) \in M^n$ be such that no x_i for $1 \leq i \leq k$ equals any of the x_j with $k < j$. Choose open neighborhoods U_{x_ℓ} of x_ℓ in M for all ℓ such that each $\overline{U_{x_i}}$ with $i \leq k$ is disjoint from any $\overline{U_{x_j}}$ with $k < j$. For smooth functions f_ℓ with support in U_{x_ℓ} for all ℓ, we have that for $i \leq k$ all functions f_i vanish on $\bigcap_{j=1}^{k}(M \setminus U_{x_j})$, thus $\mathcal{L}_{X^{\text{diag}}}(\check{G}(f_1, \ldots, f_k))|\big(\bigcap_{j=1}^{k}(M \setminus U_{x_j})\big)^{n-k} = 0$ for all $X \in \mathfrak{X}_{\text{diag}}(M, \partial M, \mu_0)$.

For $k < j$ we have $\mathrm{supp}(f_j) \subset U_{x_j} \subset \bigcap_{i=1}^{k}(M \setminus U_{x_i})$. Consider $f_1, \ldots, f_k$ as fixed. Using induction on n and replacing M by the submanifold (non-compact!) $\bigcap_{i=1}^{k}(M \setminus U_{x_i})$ we may assume that the main theorem is already true for

$$\check{G}_c(f_1, \ldots, f_k)|\big(\bigcap_{j=1}^{k}(M \setminus U_{x_j})\big)^{n-k}$$

so that

$$\begin{aligned}
\check{G}_c(f_1, \ldots, f_k)(f_{k+1}, \ldots, f_n) &= C_0(f_1, \ldots, f_k) \int f_{k+1} \ldots f_n \mu_0 \\
&+ \sum_{i=k+1}^{n} C_i(f_1, \ldots, f_k) \int_M \alpha_i \cdot \int_M f_{k+1} \ldots \widehat{f_i} \ldots f_n \, \mu_0 \\
&+ \sum_{k<i<j}^{n} C_{ij}(f_1, \ldots, f_k) \int_M f_i f_j \, \mu_0 \cdot \int_M f_{k+1} \ldots \widehat{f_i} \ldots \widehat{f_j} \ldots f_n \, \mu \\
&+ \ldots \\
&+ C_{12\ldots n}(f_1, \ldots, f_k) \int_M f_{k+1} \, \mu_0 \cdot \cdots \cdot \int_M f_n \, \mu.
\end{aligned}$$

Now all the expressions $C(f_1, \ldots, f_k)$ are again invariant, and we can subject it also to the induction hypothesis. All the resulting multilinear operators are defined on the whole of M. If we substract them from the original $\hat{G}_c$, the resulting distribution has support in the set of all $(x_1, \ldots, x_n) \in M^n$ such that $x_{i_k} = x_{j_{\ell(k)}}$ for an injective mapping $\ell : \{1, \ldots, k\} \to \{1, \ldots, n-k\}$.

Finally we end up with a distribution with support on the full diagonal $\{(x, \ldots, x) : x \in M\} \subset M^n$ whose form is determined by the last lemma. $\square$

Thank you for listening.