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

CES-Seminar
April 6, 2018,
Brno
Devoted to the memory of Thomas Friedrich
(12. Oktober 1949 in Leipzig — 27. Februar 2018 in Marburg)

Based on:

[T.Friedrich: Lecture in Winterschool in Srni, 1990 or 1991]

[T.Friedrich. Die Fisher-Information und symplektische Strukturen. Math. Nachr., 153, 273–296, 1991]

[M.Bauer, M.Bruveris, P.Michor: Uniqueness of the Fisher–Rao metric on the space of smooth densities, Bull. London Math. Soc. 48, 3 (2016), 499-506, arXiv:1411.5577]

[M.Bruveris, P.Michor: Geometry of the Fisher-Rao metric on the space of smooth densities, Mathematische Nachrichten ??, arxiv:1607.04550]

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

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 $Diff(M)$ 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 compact smooth manifolds with corners (for example, a simplex), and the full proof is given (keeping the (partial) tradition of naturality questions in CES).

The Fisher–Rao metric on the space $\mathrm{Prob}(M)$ of probability densities is of importance in the field of information geometry. Restricted to finite-dimensional submanifolds of $\mathrm{Prob}(M)$, 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.

See also [Ay, Jost, Le, Schwachhöfer: Information Geometry, 2017].

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

Manifolds with corners

A manifold with corners (recently also called a quadrantic manifold) M is a smooth manifold modelled on open subsets of $\mathbb{R}^m_{\geq 0}$. Assume it is connected and second countable; then it is paracompact and it admits smooth partitions 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, 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 $q > 0$.

The space of densities

Let M^m be a smooth manifold, possibly with corners. 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

$$\begin{tikzcd} \mathrm{Vol(M)} \arrow[r, "|\det(T\varphi^{-1})| \circ \varphi"] \arrow[d] & \mathrm{Vol(M)} \arrow[d] \\ M \arrow[r, "\varphi"] & M \end{tikzcd} .$$

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:

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

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

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

Main Theorem. [BBM, 2016] for M without boundary

Let M be a connected smooth compact manifold with corners, 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 $\Phi = (\quad)^2 : \mathrm{Dens}_+(S^1) \to \mathrm{Met}(S^1)$, $\Phi(f d\theta) = f^2 d\theta^2$.
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 many $\mathrm{Diff}(S^1)$-invariant metrics on $\mathrm{Dens}_+(S^1)$ and on $\mathrm{Prob}(S^1)$.

Moser's theorem for manifolds with corners [BMPR18]

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.

Aside: 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.]

Proof of the Main Theorem

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

$$\big((\varphi_\mu^*)^* G\big)_\mu(\alpha, \beta) = G_{\varphi_\mu^*\mu}(\varphi_\mu^*\alpha, \varphi_\mu^*\beta) = G_{c.\mu_0}(\varphi_\mu^*\alpha, \varphi_\mu^*\beta)\,.$$

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

$$G_{c\mu_0}(\alpha, \beta) = C_1(c).\int_M \frac{\alpha}{\mu_0}\frac{\beta}{\mu_0}\mu_0 + C_2(c)\int_M \alpha \cdot \int_M \beta$$

for some functions C_1, C_2 of the total volume $c = \mu(M)$. Both bilinear forms are still invariant under the action of the group $\mathrm{Diff}(M, c\mu_0) = \mathrm{Diff}(M, \mu_0) = \{\psi \in \mathrm{Diff}(M) : \psi^*\mu_0 = \mu_0\}$.

The bilinear form

$$T_{\mu_0}\,\mathrm{Dens}_+(M) \times T_{\mu_0}(M)\,\mathrm{Dens}_+ \ni (\alpha, \beta) \mapsto G_{c\mu_0}\Big(\frac{\alpha}{\mu_0}\mu_0, \frac{\beta}{\mu_0}\mu_0\Big)$$

can be viewed as a bilinear form

$$C^\infty(M) \times C^\infty(M) \ni (f, g) \mapsto G_c(f, g)\,.$$

We will consider now the associated bounded linear mapping

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

(1) The Lie algebra $\mathfrak{X}(M, \partial M, \mu_0)$ of $\mathrm{Diff}(M, \mu_0)$ consists of vector fields X which are tangent to each boundary component $\partial^q M$ with

$$0 = \mathrm{div}^{\mu_0}(X) := \frac{\mathcal{L}_X \mu_0}{\mu_0}\,.$$

On an oriented open subset $U \subset M$, each density is an m-form, $m = \dim(M)$, and $\operatorname{div}^{mu_0}(X) = d i_X \mu_0$.

The mapping $\hat{\iota}_{\mu_0} : \mathfrak{X}(U) \to \Omega^{m-1}(U)$ given by $X \mapsto i_X \mu_0$ is an isomorphism, and also

$$\begin{aligned}\hat{\iota}_{\mu_0} : \mathfrak{X}(U, \partial U) \to \Omega^{m-1}(U, \partial U) = \\ = \{\alpha \in \Omega^{m-1}(M) : j^*_{\partial^q M}\alpha = 0 \text{ for all } q \geq 1\}\end{aligned}$$

is an isomorphism onto the space of differential forms that pull back to 0 on each boundary stratum. The Lie subalgebra $\mathfrak{X}(U, \partial U, \mu_0)$ of divergence free vector fields corresponds to the space of closed $(m-1)$-forms.

Denote by $\mathfrak{X}_{\text{exact}}(M, \partial M, \mu_0)$ the set (not a vector space) of 'exact' divergence free vector fields $X = \hat{\iota}^{-1}_{\mu_0}(d\omega)$, where $\omega \in \Omega^{m-2}_c(U, \partial U)$ for an oriented open subset $U \subset M$.

(2) *If for* $f \in C^\infty(M)$ *and a connected open 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.*

Since we shall need some details later on, we prove this well-known fact. Let $x \in U \setminus \partial U$. For every tangent vector $X_x \in T_xM$ we can find a vector field $X \in \mathfrak{X}_{\text{exact}}(M, \partial M, \mu_0)$ such that $X(x) = X_x$; to see this, choose a chart (U_x, u) near x such that $U_x \subseteq U \setminus \partial U$ and $\mu_0|U_x = du^1 \wedge \cdots \wedge du^m$, and choose $g \in C_c^\infty(U_x)$, such that $g = 1$ near x. Then $X := \hat{\iota}_{\mu_0}^{-1} d(g.u^2.du^3 \wedge \cdots \wedge du^m) \in \mathfrak{X}_{\text{exact}}(M, \partial M, \mu_0)$ and $X = \partial_{u^1}$ near x. So we can produce a basis for T_xM and even a local frame near x.

Thus $\mathcal{L}_X f|U = 0$ for all $X \in \mathfrak{X}_{\text{exact}}(M, \partial M, \mu_0)$ implies $df|_{U \setminus \partial U} = 0$, thus $df = 0$ and f is constant on U.

(2∂) *Similarly, if* $x \in \partial^q M$ *and* $X_x \in T_x(\partial^q M)$ *we can find* $X \in \mathfrak{X}_{exact}(M, \partial M, \mu_0)$ *with* $X(x) = X_x$.

(3) *If for a distribution $A \in \mathcal{D}'(M)$ 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 $f \in C_c^\infty(U)$.*

Because $\langle \mathcal{L}_X A, f\rangle = -\langle A, \mathcal{L}_X f\rangle$, the invariance $\mathcal{L}_X A|U = 0$ implies $\langle A, \mathcal{L}_X f\rangle = 0$ for all $f \in C_c^\infty(U)$. Clearly, $\int_M (\mathcal{L}_X f)\mu_0 = 0$. For each $x \in U$ let $U_x \subset U$ be an open oriented chart which is diffeomorphic to $\mathbb{R}^q_{\geq 0} \times \mathbb{R}^{m-q}$. Let $g \in C_c^\infty(U_x)$ satisfy $\int_M g\mu_0 = 0$; we will show that $\langle A, g\rangle = 0$. The integral over $g\mu_0$ is zero, so the compact cohom. class $[g\mu_0] \in H_c^m(U_x, \partial U_x) \cong \mathbb{R}$ vanishes; see [BMPR2018, section 8]. Thus there exists $\alpha \in \Omega_c^{m-1}(U_x, \partial U_x) \subset \Omega^{m-1}(M, \partial M)$ with $d\alpha = g\mu_0$. Since U_x is diffeomorphic to $\mathbb{R}^q_{\geq 0} \times \mathbb{R}^{m-q}$, we can write $\alpha = \sum_j f_j d\beta_j$ with $\beta_j \in \Omega^{m-2}(U_x, \partial U_x)$ and $f_j \in C_c^\infty(U_x)$. Choose $h \in C_c^\infty(U_x)$ with $h = 1$ on $\bigcup_j \mathrm{supp}(f_j)$, so that $\alpha = \sum_j f_j d(h\beta_j)$ and $h\beta_j \in \Omega^{m-2}(M, \partial M)$. Then the vector fields $X_j = \hat{\iota}_{\mu_0}^{-1} d(h\beta_j)$ lie in $\mathfrak{X}_{\mathrm{exact}}(M, \partial M, \mu_0)$ and we have the identity $\sum_j f_j . i_{X_i}\mu_0 = \alpha$.

This means $\sum_j (\mathcal{L}_{X_j} f_j)\mu_0 = \sum_j \mathcal{L}_{X_j}(f_j\mu_0) = \sum_j di_{X_j}(f_j\mu_0) = d\Big(\sum_j f_j . i_{X_j}\mu_0\Big) = d\alpha = g\mu_0$ or $\sum_j \mathcal{L}_{X_j} f_j = g$, leading to

$$\langle A, g\rangle = \sum_j \langle A, \mathcal{L}_{X_j} f_j\rangle = -\sum_j \langle \mathcal{L}_{X_j} A, f_j\rangle = 0\,.$$

So $\langle A, g\rangle = 0$ for all $g \in C_c^\infty(U_x)$ with $\int_M g\mu_0 = 0$. Finally, choose a function φ with support in U_x and $\int_M \varphi\mu_0 = 1$. Then for any $f \in C_c^\infty(U_x)$, the function defined by $g = f - (\int_M f\mu_0).\varphi$ in $C^\infty(M)$ satisfies $\int_M g\mu_0 = 0$ and so

$$\langle A, f\rangle = \langle A, g\rangle + \langle A, \varphi\rangle \int_M f\mu_0 = C_x \int_M f\mu_0\,,$$

with $C_x = \langle A, \varphi\rangle$. Thus $A|U_x = C_x\mu_0|U_x$. Since U is connected, the constants C_x are all equal: Choose $\varphi \in C_c^\infty(U_x \cap U_y)$ with $\int \varphi\mu_0 = 1$. Thus (3) is proved.

(4) *The operator* $\check{G}_c : C^\infty(M) \to \mathcal{D}'(M)$ *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 we have* $\check{G}(f)|U = C_U(f)\mu_0|U$ *for some constant* $C_U(f)$.

For $x \in U$ choose $g \in C^\infty(M)$ with $g = 1$ near $M \setminus U$ and $g = 0$ on a neighborhood V of x. Then for any $X \in \mathfrak{X}_{\text{exact}}(M, \partial M, \mu_0)$, that is $X = \hat{\iota}_{\mu_0}^{-1}(d\omega)$ for some $\omega \in \Omega_c^{m-2}(W, \partial W)$ where $W \subset M$ is an oriented open set, let $Y = \hat{\iota}_{\mu_0}^{-1}(d(g\omega))$. The vector field $Y \in \mathfrak{X}_{\text{exact}}(M, \partial M, \mu_0)$ equals X near $M \setminus U$ and vanishes on V. Since f is constant on U, $\mathcal{L}_X f = \mathcal{L}_Y f$. For all $h \in C^\infty(M)$ we have $\left\langle \mathcal{L}_X \check{G}_c(f), h \right\rangle = \left\langle \check{G}_c(f), -\mathcal{L}_X h \right\rangle = -G_c(f, \mathcal{L}_X h) = G_c(\mathcal{L}_X f, h) = \left\langle \check{G}_c(\mathcal{L}_X f), h \right\rangle$, since G_c is invariant. Thus also

$$\mathcal{L}_X \check{G}_c(f) = \check{G}_c(\mathcal{L}_X f) = \check{G}_c(\mathcal{L}_Y f) = \mathcal{L}_Y \check{G}_c(f)\,.$$

Now Y vanishes on V and therefore so does $\mathcal{L}_X \check{G}_c(f)$. By (3) we have $\check{G}_c(f)|V = C_V(f)\mu_0|V$ for some $C_V(f) \in \mathbb{R}$. Since U is connected, all the constants $C_V(f)$ have to agree, giving a constant $C_U(f)$, depending only on U and f. Thus (4) follows.

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

$$\hat{G}_c \in \mathcal{D}'(M \times M) = (\mathcal{D}(M)\widehat{\widehat{\otimes}}\mathcal{D}(M))' = \\ = (\mathcal{D}(M)\widehat{\otimes}\mathcal{D}(M))' = L(\mathcal{D}(M), \mathcal{D}'(M)) \ni \check{G}_c$$

where one needs first the completed inductive or ϵ-tensorproduct, and then the projective one. Note the defining relations

$$G_c(f, g) = \langle \check{G}_c(f), g \rangle = \langle \hat{G}_c, f \otimes g \rangle.$$

Moreover, $\hat{G}_c$ is invariant under the diagonal action of $\mathrm{Diff}(M, \mu_0)$ on $M \times M$. In view of the tensor product in the defining relations, the infinitesimal version of this invariance is: $\mathcal{L}_{X\times 0+0\times X}\hat{G}_c = 0$ for all $X \in \mathfrak{X}(M, \partial M, \mu_0)$.

(5) *There exists a constant* $C_2 = C_2(c)$ *such that the distribution* $\hat{G}_c - C_2\mu_0 \otimes \mu_0$ *is supported on the diagonal of* $M \times M$.

Namely, if $(x, y) \in M \times M$ is not on the diagonal, then there exist open neighborhoods U_x of x and U_y of y in M such that $\overline{U_x} \times \overline{U_y}$ is disjoint to the diagonal, or $\overline{U_x} \cap \overline{U_y} = \emptyset$. Choose any functions $f, g \in C^\infty(M)$ with $\operatorname{supp}(f) \subset U_x$ and $\operatorname{supp}(g) \subset U_y$. Then $f|(M \setminus \overline{U_x}) = 0$, so by (4), $\check{G}_c(f)|(M \setminus \overline{U_x}) = C_{M\setminus\overline{U_x}}(f).\mu_0$. Therefore,

$$
\begin{aligned}
G_c(f, g) &= \langle \hat{G}_c, f \otimes g\rangle = \langle \check{G}_c(f), g\rangle \\
&= \langle \check{G}_c(f)|(M \setminus \overline{U_x}), g|(M \setminus \overline{U_x})\rangle\,, \quad \text{since } \operatorname{supp}(g) \subset U_y \subset M \setminus \overline{U_x}, \\
&= C_{M\setminus\overline{U_x}}(f) \cdot \int_M g\mu_0
\end{aligned}
$$

By applying the argument for the transposed bilinear form $G_c^T(g, f) = G_c(f, g)$, which is also $\operatorname{Diff}(M, \mu_0)$-invariant, we get

$$
G_c(f, g) = G_c^T(g, f) = C'_{M\setminus\overline{U_y}}(g) \cdot \int_M f\mu_0\,.
$$

Fix two functions f_0, g_0 with the same properties as f, g and additionally $\int_M f_0\mu_0 = 1$ and $\int_M g_0\mu_0 = 1$. Then we get $C_{M\setminus\overline{U_x}}(f) = C'_{M\setminus\overline{U_y}}(g_0)\int_M f\mu_0$, and so

$$\begin{aligned} G_c(f,g) &= C'_{M\setminus\overline{U_y}}(g_0)\int_M f\mu_0 \cdot \int_M g\mu_0 \\ &= C_{M\setminus\overline{U_x}}(f_0)\int_M f\mu_0 \cdot \int_M g\mu_0 \,. \end{aligned}$$

Since $\dim(M) \geq 2$ and M is connected, the complement of the diagonal in $M \times M$ is also connected, and thus the constants $C_{M\setminus\overline{U_x}}(f_0)$ and $C'_{M\setminus\overline{U_y}}(g_0)$ cannot depend on the functions f_0, g_0 or the open sets U_x and U_y as long as the latter are disjoint. Thus there exists a constant $C_2(c)$ such that for all $f, g \in C^\infty(M)$ with disjoint supports we have

$$G_c(f,g) = C_2(c)\int_M f\mu_0 \cdot \int_M g\mu_0$$

Since $C_c^\infty(U_x \times U_y) = C_c^\infty(U_x)\bar{\otimes} C_c^\infty(U_y)$, this implies claim (5).

Now we can finish the proof. We may replace $\hat{G}_c \in \mathcal{D}'(M \times M)$ by $\hat{G}_c - C_2\mu_0 \otimes \mu_0$ and thus assume without loss that the constant C_2 in (5) is 0. Let (U, u) be an oriented chart on M such that $\mu_0|U = du^1 \wedge \cdots \wedge du^m$, and let $\tilde{U}$ be an extension of U to a smooth manifold without boundary with an extension of the chart mapping u. The distribution $\hat{G}_c|U \times U \in \mathcal{D}'(U \times U) \subset \mathcal{D}'(\tilde{U} \times \tilde{U})$ has support contained in the diagonal and is of finite order k. By [Hörmander I, 1983, Theorem 5.2.3], the corresponding operator $\check{G}_c : C_c^\infty(U) \to \mathcal{D}'(U)$ is of the form $\hat{G}_c(f) = \sum_{|\alpha| \le k} A_\alpha.\partial^\alpha f$ for $A_\alpha \in \mathcal{D}'(U)$, so that $G(f, g) = \langle \check{G}_c(f), g\rangle = \sum_\alpha \langle A_\alpha, (\partial^\alpha f).g\rangle$. Moreover, the A_α in this representation are uniquely given, as is seen by a look at [Hörmander I, 1983, Theorem 2.3.5].

For $x \in U$ choose an open set U_x with $x \in U_x \subset \overline{U_x} \subset U$, and choose $X \in \mathfrak{X}_{\text{exact}}(M, \partial M, \mu_0)$ with $X|U_x = \partial_{u^i}$ (tangential to the boundary), as in (2∂). For functions $f, g \in C_c^\infty(U_x)$ we then have, by the invariance of G_c,

$$\begin{aligned}0 &= G_c(\mathcal{L}_X f, g) + G_c(f, \mathcal{L}_X g) = \langle \hat{G}_c | U \times U, \mathcal{L}_X f \otimes g + f \otimes \mathcal{L}_X g \rangle \\ &= \sum_\alpha \langle A_\alpha, (\partial^\alpha \partial_{u^i} f).g + (\partial^\alpha f)(\partial_{u^i} g) \rangle \\ &= \sum_\alpha \langle A_\alpha, \partial_{u^i}((\partial^\alpha f).g) \rangle = \sum_\alpha \langle -\partial_{u^i} A_\alpha, (\partial^\alpha f).g \rangle .\end{aligned}$$

Since the corresponding operator has again a kernel distribution which is supported on the diagonal, and since the distributions in the representation are unique, we can conclude that $\partial_{u^i} A_\alpha | U_x = 0$ for each α, and each i such that ∂_{u^i} is tangential to the boundary.

To see that this implies that $A_\alpha|U_x = C_\alpha \mu_0|U_x$, let $f \in C_c^\infty(U_x)$ with $\int_M f\mu_0 = 0$. Then, as in (3), there exists $\omega \in \Omega_c^{m-1}(U_x, \partial U_x)$ with $d\omega = f\mu_0$. We have $\omega = \sum_i \omega_i . du^1 \wedge \cdots \wedge \widehat{du^i} \wedge du^m$ (only those i with ∂_{u^i} tangential to the boundary have $\omega_i \neq 0$), and so $f = \sum_i (-1)^{i+1} \partial_{u^i} \omega_i$ with $\omega_i \in C_c^\infty(U_x)$. Thus

$$\langle A_\alpha, f\rangle = \sum_i (-1)^{i+1} \langle A_\alpha, \partial_{u^i}\omega_i\rangle = \sum_i (-1)^i \langle \partial_{u^i} A_\alpha, \omega_i\rangle = 0\,.$$

Hence $\langle A_\alpha, f\rangle = 0$ for all $f \in C_c^\infty(U_x)$ with zero integral and as in the proof of (3) we can conclude that $A_\alpha|U_x = C_\alpha \mu_0|U_x$.

But then $G_c(f,g) = \int_{U_x}(Lf).g\mu_0$ for the differential operator $L = \sum_{|\alpha|\le k} C_\alpha \partial^\alpha$ with constant coefficients on U_x. Now we choose $g \in C_c^\infty(U_x)$ such that $g = 1$ on the support of f. By the invariance of G_c we have again

$$0 = G_c(\mathcal{L}_X f, g) + G_c(f, \mathcal{L}_X g) = \int_{U_x} L(\mathcal{L}_X f).g\mu_0 + \int_{U_x} L(f).\mathcal{L}_X g.\mu_0$$
$$= \int_{U_x} L(\mathcal{L}_X f)\mu_0 + 0$$

for each $X \in \mathfrak{X}(M, \partial M, \mu_0)$. Thus the distribution $f \mapsto \int_{U_x} L(f)\mu_0$ vanishes on all functions of the form $\mathcal{L}_X f$, and by (3) we conclude that $L(\quad).\mu_0 = C_x.\mu_0$ in $\mathcal{D}'(U_x)$, or $L = C_x \,\mathrm{Id}$. By covering M with open sets U_x, we see that all the constants C_x are the same. This concludes the proof of the Main Theorem. $\square$

Thank you for listening.