

This is a postprint version of the following published document:

Ciaglia, F. M., Marmo, G. y Pérez Pardo, J.M. (2018). Generalized potential functions in differential geometry and information geometry. *International Journal of Geometric Methods in Modern Physics*, 16(01), 1940002.

DOI: <https://doi.org/10.1142/S0219887819400024>

# Generalised Potential Functions in Differential Geometry and Information Geometry

F.M. Ciaglia<sup>1,2</sup>, G. Marmo<sup>1,2</sup>, and J.M. Pérez-Pardo<sup>3,4</sup>

<sup>1</sup>*Dipartimento di Fisica “E. Pancini”, Università di Napoli Federico II, Complesso Universitario di Monte S. Angelo Edificio 6, via Cintia, 80126 Napoli, Italy.*

<sup>2</sup>*INFN-Sezione di Napoli, Complesso Universitario di Monte S. Angelo Edificio 6, via Cintia, 80126 Napoli, Italy.*

<sup>3</sup>*Depto. de Matemáticas, Univ. Carlos III de Madrid, Avda. de la Universidad 30, 28911 Leganés, Madrid, Spain.*

<sup>3</sup>*Instituto de Ciencias Matemáticas (CSIC-UAM-UC3M-UCM), C./ Nicolás Cabrera 13-15. Campus de Cantoblanco, UAM, 28049 Madrid, Spain.*

April 27, 2018

Subject Classification 53C99, 53Z05, 53A40

Keywords: *Information Geometry, Generalised Potential Functions, Inverse Problem*

## Abstract

Potential functions can be used as generating potentials of relevant geometric structures for a Riemannian manifold such as the Riemannian metric and affine connections. We study whether this procedure can also be applied to tensors of rank four and find a negative answer. We study this from the perspective of solving the inverse problem and also from an intrinsic point of view.

## 1 Introduction

In 1925 R.A. Fisher [18] introduced the celebrated concept of statistical distance. This notion was the cornerstone for what has developed nowadays into a major field in Statistics, known as Information Geometry. This has ramifications, for instance, in Information Theory, Statistical Mechanics, Biology and Quantum Mechanics, see [2,8]. One key ingredient of the successful approach of Information Geometry has been the geometrisation. That is, aiming at stating the results in a way that is independent of the specific situation or coordinates and thus providing a very general framework that can be applied to a variety of situations.

The central object of the approach is known as a statistical manifold and we will denote it by  $\mathcal{S}$ . This assumes that, given a particular family of probability distributions  $\mathcal{P}(X)$  on a measure space  $X$ , one is able to find a parametrisation,

i.e. a mapping

$$p : \Xi \rightarrow \mathcal{P}(X),$$

such that  $p$  is a bijection onto its image  $p(\Xi)$ . Under suitable conditions, we refer to [2] for a proper definition of a Statistical Manifold,  $\mathcal{S}$  has the structure of a differentiable manifold and Fisher's statistical distance provides it with a natural Riemannian metric, the so called Fisher-Rao metric [27].

In addition to this Riemannian structure, there are two geometric objects related to it that play an important role. On one hand there is a so called dualistic structure [1]. This is a pair of torsion-free connections  $(\nabla, \nabla^*)$  that satisfy the relation

$$X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X^* Z),$$

where  $X, Y, Z$  are vector fields over the manifold. These two connections are said to be dual to each other. Notice that, in particular, the Levi-Civita connection is the unique self-dual connection. On the other hand there is a notion of Divergence function, the most notable example of which is commonly known as Kullback-Leibler divergence (also called relative entropy) [22]. A divergence function is a two-point function

$$F : \mathcal{S} \times \mathcal{S} \rightarrow \mathbb{R}$$

which is differentiable in both entries and with the following properties:

$$\begin{aligned} F(\mathbf{x}, \mathbf{y}) &\geq 0. \\ F(\mathbf{x}, \mathbf{y}) &= 0 \Leftrightarrow \mathbf{x} = \mathbf{y}. \\ \frac{\partial F}{\partial x^i} \Big|_{\mathbf{x}=\mathbf{y}} &= \frac{\partial F}{\partial y^j} \Big|_{\mathbf{x}=\mathbf{y}} = 0, \end{aligned} \tag{1.1}$$

for  $i, j = 1, \dots, \dim(\mathcal{S})$ , and where  $\mathbf{x} = \{x^1, x^2, \dots\}$  and  $\mathbf{y} = \{y^1, y^2, \dots\}$  are coordinates of the manifold  $\mathcal{S}$ . The last condition just requires that the first derivatives of the function  $F$  vanish on the diagonal. Notice that the three conditions together amount to require that the function  $F$  has a local extremum on the diagonal.

In the context of information geometry and of statistical inference the divergence functions are measures of the relative difference between probability distributions. We refer to [2] for an introduction and selected applications of divergence functions in Information Geometry. One remarkable property, from the geometric viewpoint, of these general divergence functions is that they are *generating potentials* for the geometric structures introduced so far. The concrete meaning of the term "generating potential" will be made clearer in the next section. The properties of Eq. (1.1) are sufficient for a two-point function to generate a Riemannian metric and a dualistic structure and therefore can be applied in general for any differential manifold, regardless if it is a Statistical Manifold or not. The particular case of the Kullback-Leibler divergence gives rise to the Fisher-Rao metric and the so called exponential connection and mixture connections.

Having a generating potential for geometric quantities is an advantage, for instance, when deriving geometric or, more properly, tensorial quantities that must be preserved under concrete transformations. Indeed, this is the case of

the Kullback-Leibler divergence, where its monotonicity under transformations of the probability distributions reflects on the monotonicity of the Fisher-Rao metric [12, 13, 29].

In the context of Quantum Mechanics, the geometric description of the space of pure states [3, 15] leads to a natural interpretation of the Fubini-Study metric as a quantum counterpart of the Fisher-Rao metric [17, 19, 30]. However, when one wants to apply the same information geometric approach to the space of quantum mixed states the picture is not so clear. There are intrinsic difficulties in the problem, like the ambiguity in defining the logarithmic derivative of a matrix [16, 21] and the fact that the space of mixed states is not a manifold but it is partitioned into the disjoint union of differential manifolds labelled by the rank of quantum states [9, 20, 24]. This leads to a wide variety of quantum divergence functions and quantum Fisher-Rao metrics depending on the particular aspect that one wants to highlight, cf. [11, 26]. Having a geometric description of the structures involved, independent of the particular realisation or setting, becomes useful in identifying the most relevant among all the possibilities.

The aim of this article is to analyse thoroughly the role of two-point functions in differential geometry as potential functions for tensorial quantities and we will do so in an intrinsic approach. Our interest relies also in the definition of a well-posed inverse problem for the potential function, cf. [7, 10]. In particular we will see that Hamilton's principal function is a particular solution to the problem.

## 2 Generating Potentials for Tensorial Quantities

Generating potentials are common objects in mathematics. The most simple intrinsic example that one might think of would be the exterior differential of a smooth function  $f$  on a differentiable manifold  $\mathcal{M}$ . The one-form obtained this way is intrinsically defined and does not depend on the particular coordinates chosen to define it. At this point one can also recognise that there might be some global problems when defining the inverse problem. Poincaré's Lemma shows that the inverse problem, i.e., if given a closed form  $\alpha$  one is able to find a function on  $\mathcal{M}$  such that  $df = \alpha$ , has a local solution but in general a global solution might not exist.

An interesting situation is encountered in the context of Kähler manifolds where one can describe (at least locally) the symplectic form  $\omega$  using the so-called (local) Kähler potential  $f$  by means of the formula

$$\omega_{ij} = \frac{\partial^2 f}{\partial z \partial \bar{z}}, \quad (2.1)$$

where  $z, \bar{z}$  are, respectively, holomorphic and anti-holomorphic coordinates. Again, this definition is intrinsic and the symplectic form defined this way does not depend on the particular coordinates chosen to define it. Of course, by formula (2.1) one computes the coordinate expression of the symplectic form, but this coordinate expression transforms properly under a change of coordinates of the manifold. In fact, this situation is analogous to that of the exterior differential. One can show that a local Kähler potential always exist. One can easily define the coordinates of a tensor by taking iterated derivatives of a function

defined on the manifold. For instance, in [14] there is introduced the concept of *Hessian Riemannian structure*. Let  $\mathbf{x} = \{x^1, x^2, \dots\}$  be local coordinates on an open neighbourhood  $U \subset \mathcal{M}$  and let  $f : U \rightarrow \mathbb{R}$  be a convex function. Then one can define the local coordinates of a Riemannian metric by

$$g_{ij}(\mathbf{x}) := \frac{\partial^2 f}{\partial x^i \partial x^j}.$$

While this can be done to define the local coordinates of a tensor, if one chooses to express the function  $f$  in a different local chart  $\mathbf{y} = \{y^1, y^2, \dots\}$ , the expression

$$\tilde{g}_{ij}(\mathbf{y}) := \frac{\partial^2 f}{\partial y^i \partial y^j}$$

defines a different tensor. There is a notable exception, closely related to Kählerian functions, the so-called (flat) *Hessian Structure* [28]. Hence, recovering desired transformation properties of  $g$  by imposing invariance conditions on  $f$  is not possible by such an approach. Remarkably, if one considers two-point functions with conditions (1.1) one is able to derive tensors of rank higher than one. This is well known for the information geometry community, where the relevant divergence functions that arise give rise to the Fisher-Rao metric. However, it turns out to be a general situation that can be defined on an arbitrary differentiable manifold.

Since we want to define the potential functions in an intrinsic way, we need to define the restriction to the diagonal appropriately. Consider the diagonal embedding given by

$$\begin{aligned} D : \mathcal{M} &\hookrightarrow \mathcal{M} \times \mathcal{M} \\ p &\mapsto (p, p) \end{aligned} .$$

The restriction to the diagonal is given by means of the pull-back by  $D$ . If  $F : \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}$  is a two-point function, then  $F|_{\text{diag}} = D^*F$ . In local coordinates one has that

$$D^*F(p) = F(D(p)) = F(p(x^1, x^2, \dots), p(x^1, x^2, \dots)), p \in \mathcal{M}.$$

In addition, one can define the two canonical projections:

$$\begin{aligned} \pi_L : \mathcal{M} \times \mathcal{M} &\hookrightarrow \mathcal{M} & \pi_R : \mathcal{M} \times \mathcal{M} &\hookrightarrow \mathcal{M} \\ (p, q) &\mapsto p & (p, q) &\mapsto q \end{aligned} .$$

By means of these one can define the Left-lift and Right-lift of a vector field  $X \in \mathfrak{X}(\mathcal{M})$  as follows [11]. The Left-lift  $X_L \in \mathfrak{X}(\mathcal{M} \times \mathcal{M})$  of the vector field  $X \in \mathfrak{X}(\mathcal{M})$  is defined as the unique  $\pi_L$ -related vector field such that  $X_L(\pi_R^*f) = 0$  and  $X_L(\pi_L^*f) = \pi_L^*(X(f))$ . The Right-lift is defined analogously.

We will say that a two-point function

$$S : \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}$$

is a potential function if it satisfies the property

$$D^*(\mathcal{L}_X S) = 0, \quad X \in \mathfrak{X}(\mathcal{M} \times \mathcal{M}), \quad (2.2)$$

where  $\mathcal{L}_X$  is the Lie derivative with respect to the vector field  $X$ . The condition in (2.2) is slightly more general than the conditions in (1.1). Indeed, the latter imply that the diagonal is a local extremum of the function while (2.2) only requires it to be a critical point.

Any potential function defines a second order tensor in the following way:

$$g(X, Y) := D^*(\mathcal{L}_{X_L} \mathcal{L}_{Y_L} S).$$

We will show that  $g$  is  $f$ -linear in both entries and therefore that it is a tensor. First notice that for any function  $f$  on  $\mathcal{M}$  and any vector field  $X \in \mathfrak{X}(\mathcal{M})$  one has that

$$(fX)_L = (\pi_L^* f) X_L$$

and therefore  $f$ -linearity in the first entry follows by the  $f$ -linearity of the Lie derivative. Now consider

$$\mathcal{L}_{X_L} \mathcal{L}_{(fY)_L} S = \mathcal{L}_{X_L} (\pi_L^* f \mathcal{L}_{Y_L} S) \quad (2.3)$$

$$= \pi_L^* f \mathcal{L}_{X_L} \mathcal{L}_{Y_L} S + \mathcal{L}_{X_L} (\pi_L^* f) \mathcal{L}_{Y_L} S, \quad (2.4)$$

and taking the pull-back by  $D$  on both sides we get

$$g(X, fY) = fg(X, Y) + D^*(\mathcal{L}_{X_L} (\pi_L^* f)) D^*(\mathcal{L}_{Y_L} S) \quad (2.5)$$

and by condition (2.2) the second term at the right hand side vanishes. It is easy to see that this tensor is symmetric. Indeed,

$$\mathcal{L}_{X_L} \mathcal{L}_{Y_L} S = \mathcal{L}_{Y_L} \mathcal{L}_{X_L} S + \mathcal{L}_{[X, Y]} S$$

and after applying  $D^*$  one gets

$$g(X, Y) = g(Y, X).$$

Notice that no assumption has been made on the symmetry of  $S$  under permutation of its entries and in the most common situations this will not be the case. There are other possible combinations in order to obtain tensors of rank two, but all of them lead to the same symmetric tensor up to a sign. Indeed,

$$\begin{aligned} D^*(\mathcal{L}_{X_L} \mathcal{L}_{Y_L} S) &= D^*(\mathcal{L}_{X_R} \mathcal{L}_{Y_R} S) \\ &= -D^*(\mathcal{L}_{X_R} \mathcal{L}_{Y_L} S) \\ &= -D^*(\mathcal{L}_{X_L} \mathcal{L}_{Y_R} S) = g(X, Y). \end{aligned} \quad (2.6)$$

A general and coordinate free proof of this fact can be found in [11]. Note that if every point on the diagonal of  $\mathcal{M} \times \mathcal{M}$  is a minimum for  $S$ , then  $g$  is positive semi-definite. Analogously, if every point on the diagonal of  $\mathcal{M} \times \mathcal{M}$  is a maximum for  $S$ , then  $g$  is negative semi-definite. If  $\{x^j, y^k\}$  is a coordinate system on  $\mathcal{M} \times \mathcal{M}$  adapted to the Cartesian product structure, then the local components of  $g$  read:

$$g_{jk} = \left( \frac{\partial^2 S}{\partial x^j \partial x^k} \right)_{x=y} = \left( \frac{\partial^2 S}{\partial y^j \partial y^k} \right)_{x=y} = - \left( \frac{\partial^2 S}{\partial x^j \partial y^k} \right)_{x=y}, \quad (2.7)$$

in complete accordance with the expressions used in information geometry [1, 2].

*Example 1.* Let  $\mathcal{H}$  be a finite-dimensional complex Hilbert space, and consider the open submanifold  $\mathcal{H}_0$  consisting of  $\mathcal{H}$  without the null vector. It is well known that the space of pure quantum states may be identified with the complex projective space  $\mathbb{C}\mathbb{P}(\mathcal{H})$  associated with  $\mathcal{H}$ . Furthermore, the complex projective space is the base space of the principal fibre bundle  $\pi: \mathcal{H}_0 \rightarrow \mathbb{C}\mathbb{P}(\mathcal{H})$  with structure group  $\mathbb{C}_0$ , that is, the Abelian multiplicative group of non-zero complex numbers. In particular, we have  $\pi(\psi) = [\psi]$ , where  $[\psi]$  is the equivalence class of vectors in  $\mathcal{H}_0$  differing only by the multiplication with a non-zero complex number. Now, let us consider the following two-point function on  $\mathcal{H}_0 \times \mathcal{H}_0$ :

$$S(\psi, \phi) := \frac{|\langle \psi | \phi \rangle|^2}{|\psi|^2 |\phi|^2}, \quad (2.8)$$

which is the quantum-mechanical counterpart of the (non-normalised) “generalised transition probability” introduced by Cantoni in [5, 6, 30]. If  $\{\mathbf{e}_j\}_{1, \dots, \dim(\mathcal{H})}$  is an orthonormal basis in  $\mathcal{H}$ , we may introduce a Cartesian coordinate system on  $\mathcal{H}_0$  setting  $\psi = \sum_j (x^j + iy^j) \mathbf{e}_j$ . Then, in an obvious way, we may introduce a Cartesian coordinate system  $\{x^j, y^j; X^k, Y^k\}$  on  $\mathcal{H}_0 \times \mathcal{H}_0$  that is adapted to the product structure. In this (global) coordinate system, it is a matter of straightforward computations to prove that  $S$  is a potential function. Furthermore, from the expressions:

$$\begin{aligned} \frac{\partial S}{\partial x^j} &= \frac{1}{|\psi|^2 |\phi|^2} \left( \langle \psi | \phi \rangle (X^j - iY^j) + \langle \phi | \psi \rangle (X^j + iY^j) - \frac{2x^j |\langle \psi | \phi \rangle|^2}{|\psi|^2} \right), \\ \frac{\partial S}{\partial y^j} &= \frac{1}{|\psi|^2 |\phi|^2} \left( \langle \psi | \phi \rangle (Y^j + iX^j) + \langle \phi | \psi \rangle (Y^j - iX^j) - \frac{2y^j |\langle \psi | \phi \rangle|^2}{|\psi|^2} \right), \end{aligned}$$

it is easy to see that the symmetric covariant tensor  $g$  generated by  $S$  reads:

$$\begin{aligned} g &= \sum_{j,k} \frac{2(x^j x^k + y^j y^k - \delta_{jk} R^2)}{R^4} (dx^j \otimes dx^k + dy^j \otimes dy^k) + \\ &+ \sum_{j,k} \frac{y^j x^k - y^k x^j}{R^4} (dx^j \otimes dy^k - dy^j \otimes dx^k), \end{aligned} \quad (2.9)$$

where  $R^2 = \mathbf{x}^2 + \mathbf{y}^2$ . This is the pull-back to  $\mathcal{H}_0$  of the Fubini-Study metric tensor on the space of pure quantum states  $\mathbb{C}\mathbb{P}(\mathcal{H})$ .

We may take a further step and give a generating algorithm for covariant tensors of order three. The main ingredient in order to do so is that one can combine the action of Left-lifts and Right-lifts in order to cancel non-tensorial contributions of the higher order derivatives. Consider for example

$$\begin{aligned} T_1(X, Y, Z) &:= D^* (\mathcal{L}_{X_L} \mathcal{L}_{Y_R} \mathcal{L}_{Z_R} S - \mathcal{L}_{X_R} \mathcal{L}_{Y_L} \mathcal{L}_{Z_L} S), \\ T_2(X, Y, Z) &:= D^* (\mathcal{L}_{X_L} \mathcal{L}_{Y_L} \mathcal{L}_{Z_R} S - \mathcal{L}_{X_R} \mathcal{L}_{Y_R} \mathcal{L}_{Z_L} S). \end{aligned} \quad (2.10)$$

It is easy to check that both objects are indeed  $f$ -linear in all the entries and symmetric. One can define up to 8 different possible combinations that lead to order three tensors. As happens for the second order tensor, it can be proven that all of them are equal up to a sign, cf. [11].

In the context of information geometry these tensors are called the skewness tensors [1, Chapter 4] and they are used to define the dualistic structure of the statistical manifold. The Christoffel symbols of the dual connections  $\nabla$  and  $\nabla^*$  are defined by adding and subtracting, respectively, to the Christoffel symbol of the Levi-Civita connection the skewness tensor

$$\Gamma_{\pm jk}^i = \Gamma_{jk}^i \pm g^{il} T_{ljk}.$$

For instance, by this procedure the Kullback-Leibler divergence gives rise to the exponential connection and the mixture connection. Notice that both right hand sides in (2.10) vanish if the potential function is symmetric, thus recovering the well known result that symmetric divergence functions can lead only to the self-dual case  $\nabla = \nabla^*$ .

At this point, a natural question would be if it were possible to use the same procedure to recover higher order tensors from a given potential function  $S$ . However, before considering this issue in Section 4, we will address first the inverse problem of finding a potential function that generates two given covariant tensor fields of order two and three.

### 3 Hamilton-Jacobi approach to Divergence Functions and the inverse problem

In this section we will address the inverse problem of finding a potential function given a Riemannian metric and a skewness tensor. This inverse problem will be relevant in situations where there are not clear candidates for the potential/divergence functions but where there are relevant metrics or connections that one would want to generate. This is of particular importance in Quantum Mechanics where, as explained in the introduction, it can help to distinguish or highlight particular families of divergence functions among the available variety.

In some situations one can find a canonical divergence function. For instance, the self-dual case has as natural divergence function the geodesic distance between points of the manifold. In general, a canonical solution to the problem will not exist. Roughly speaking, this is due to the fact that the metric and skewness tensor only fix the lower order terms in a power series expansion. Thus, the higher order terms might be changed arbitrarily to define infinite possible *equivalent* divergence functions. By equivalent we mean that they generate the same metric and skewness tensors. A successful attempt to obtain solutions to the inverse problem is given, for instance, in [4], however, this approach lacks of a characterisation of the geometric conditions needed to guarantee the global existence of the solution.

We will review here the approach introduced in [10] based on the Hamilton-Jacobi theory. In this case the existence of the solution of the inverse problem can be characterised. Moreover, it will allow us to address the problem of finding a generating potential for higher order rank tensors. The main idea is to consider Hamilton's Principal function as a two-point function. Given a Lagrangian  $\mathfrak{L}$  and a solution of the equations of motion determined by the Lagrangian,  $\gamma : [t_{\text{in}}, t_{\text{fin}}] \rightarrow \mathcal{M}$  Hamilton's principal function is given by the action integral

$$S[\gamma] = \int_{t_{\text{in}}}^{t_{\text{fin}}} \mathfrak{L}(\gamma(t), \dot{\gamma}(t)) dt.$$

If one considers only solutions with  $t_{\text{in}} = 0$  and  $t_{\text{fin}} = 1$  such that  $\gamma(0) = \mathbf{x}$  and  $\gamma(1) = \mathbf{y}$  it becomes a two point function

$$S[\mathbf{x}, \mathbf{y}] = \int_0^1 \mathfrak{L}(\gamma(t), \dot{\gamma}(t)) dt.$$

Moreover, assume that the dynamical system defined by the Lagrangian is completely integrable.

From now on we will assume that the Lagrangian is at least of second order with respect to the velocities, i.e.,

$$\mathfrak{L} = g_{ij}v^i v^j + \mathcal{O}(|v|^3),$$

and that  $g$  is positive defined. In particular, let us consider the Lagrangian

$$\mathfrak{L}(\mathbf{q}, \mathbf{v}) = \frac{1}{2}g_{ij}v^i v^j + \frac{\alpha}{6}T_{ijk}v^i v^j v^k, \quad (3.1)$$

where  $\alpha \in \mathbb{R}$  is a parameter that is included for convenience. Have into account that this Lagrangian can be written in a coordinate free way as was done in [10].

We will show that the two-point function  $S$  determined by this Lagrangian solves, under certain assumptions, the inverse problem for  $g$  and  $T$ . From the Hamilton-Jacobi theory [23], having into account that  $S$  is Hamilton's principal function and recalling that in the Lagrangian formulation of dynamics [25] the local expression of the canonical momentum is given by  $p_i = \frac{\partial \mathfrak{L}}{\partial v^i}$ , we have that:

$$\frac{\partial S}{\partial x^i} = -p_i = \left. \frac{\partial \mathfrak{L}}{\partial v^i} \right|_{\mathbf{q}=\mathbf{x}} = -g_{ij}(\mathbf{x})v_{\text{in}}^j - \frac{\alpha}{2}T_{ijk}(\mathbf{x})v_{\text{in}}^j v_{\text{in}}^k, \quad (3.2)$$

$$\frac{\partial S}{\partial y^i} = P_i = \left. \frac{\partial \mathfrak{L}}{\partial v^i} \right|_{\mathbf{q}=\mathbf{y}} = g_{ij}(\mathbf{y})v_{\text{fin}}^j + \frac{\alpha}{2}T_{ijk}(\mathbf{y})v_{\text{fin}}^j v_{\text{fin}}^k. \quad (3.3)$$

This guarantees that

$$D^* p_i = D^* P_i = 0, \quad i = 1, \dots, n.$$

This is because the points in the diagonal satisfy  $\mathbf{x} = \gamma(0) = \gamma(1) = \mathbf{y}$  and therefore the initial velocity is zero. This shows that the two-point function satisfies the condition in Eq. (2.2).

We need to obtain the explicit solution of the equations of motion in order to be able to express  $\mathbf{v}_{\text{in}}$  and  $\mathbf{v}_{\text{fin}}$  only in terms of the endpoints of the curve  $\gamma$ . If we use the Taylor expansion around  $t = 0$  of the solution of the equations of motion we have that

$$\gamma(t) = \gamma(0) + t\mathbf{v}_{\text{in}} + \frac{t^2}{2}\dot{\mathbf{v}}_{\text{in}} + \frac{t^3}{6}\ddot{\mathbf{v}}_{\text{in}} + \mathcal{O}(t^4). \quad (3.4)$$

Substituting  $t = 1$  and having into account that  $\gamma(0) = \mathbf{x}$  and  $\gamma(1) = \mathbf{y}$  we have that

$$\mathbf{v}_{\text{in}} = \mathbf{y} - \mathbf{x} - \frac{1}{2}\dot{\mathbf{v}}_{\text{in}} - \frac{1}{6}\ddot{\mathbf{v}}_{\text{in}} + \mathcal{O}(t^4). \quad (3.5)$$

Analogously, if we expand around  $t = 1$  we get an expression for  $\mathbf{v}_{\text{fin}}$ , namely,

$$\mathbf{v}_{\text{fin}} = \mathbf{y} - \mathbf{x} + \frac{1}{2}\dot{\mathbf{v}}_{\text{fin}} - \frac{1}{6}\ddot{\mathbf{v}}_{\text{fin}} + \mathcal{O}(t^4). \quad (3.6)$$

The Euler-Lagrange equations determined by the Lagrangian are

$$\dot{v}^l = -\alpha T_{jk}^l v^j v^k - \Gamma_{jk}^l v^j v^k - \frac{\alpha}{6} g^{lr} A_{rjks} v^j v^k v^s, \quad (3.7)$$

where we are raising and lowering indices in the usual way by means of the Riemannian metric  $g$ . Here,  $\Gamma_{jk}^i$  are the Christoffel symbols of the Levi-Civita connection and we have defined the tensor

$$A_{rjks} := \left( \frac{\partial T_{jkr}}{\partial q^s} + \frac{\partial T_{jrs}}{\partial q^k} + \frac{\partial T_{rks}}{\partial q^j} - \frac{\partial T_{jks}}{\partial q^r} \right).$$

We will assume now that  $\dot{\mathbf{v}}$  is an analytic function of the velocities and we will neglect terms with high powers of  $\mathbf{v}$ , since we are interested in what happens in a neighbourhood of  $\mathbf{v} = 0$ , i.e. the diagonal. We can inductively substitute (3.7) in (3.5) and (3.6) in order to obtain the expressions for  $\mathbf{v}_{\text{in}}$  and  $\mathbf{v}_{\text{fin}}$  respectively. The analytic expression for  $\dot{\mathbf{v}}$  is obtained by differentiating the equations of motion. Notice that  $\dot{\mathbf{v}}$  is at least of order  $\mathcal{O}(\mathbf{v}^2)$  and that  $\ddot{\mathbf{v}}$  is at least of order  $\mathcal{O}(\mathbf{v}^3)$ . If we define now  $\Delta^i = y^i - x^i$  a straightforward calculation leads to:

$$\begin{aligned} \frac{\partial S}{\partial x^i} = & -g_{ij}(\mathbf{x})\Delta^j - \frac{1}{2}\Gamma_{ijk}(\mathbf{x})\Delta^j\Delta^k - \frac{1}{6}\Gamma_{ijk}(\mathbf{x})\Gamma_{ls}^j(\mathbf{x})\Delta^k\Delta^l\Delta^s \\ & - \frac{\alpha}{2}T_{ijk}(\mathbf{x})\Delta^j\Delta^k - \frac{\alpha}{12}A_{ijkl}(\mathbf{x})\Delta^j\Delta^k\Delta^l + \mathcal{O}((\mathbf{y} - \mathbf{x})^4). \end{aligned} \quad (3.8)$$

$$\begin{aligned} \frac{\partial S}{\partial y^i} = & g_{ij}(\mathbf{y})\Delta^j - \frac{1}{2}\Gamma_{ijk}(\mathbf{y})\Delta^j\Delta^k + \frac{1}{6}\Gamma_{ijk}(\mathbf{y})\Gamma_{ls}^j(\mathbf{y})\Delta^k\Delta^l\Delta^s \\ & + \frac{\alpha}{2}T_{ijk}(\mathbf{y})\Delta^j\Delta^k - \frac{\alpha}{12}A_{ijkl}(\mathbf{y})\Delta^j\Delta^k\Delta^l + \mathcal{O}((\mathbf{y} - \mathbf{x})^4). \end{aligned} \quad (3.9)$$

Now one can check that formulae (2.6) and (2.10) give rise to the tensors  $g$  and  $T$ . For instance, one has that

$$D^* \left( \frac{\partial^2 S}{\partial y^i \partial x^j} \right) (\mathbf{q}) = D^* \left( \frac{\partial^2 S}{\partial x^i \partial y^j} \right) (\mathbf{q}) = -g_{ij} = -D^* \left( \frac{\partial^2 S}{\partial x^i \partial x^j} \right) (\mathbf{q}).$$

this shows that (up to a sign) only one tensor of rank two can be generated out of the two-point function. Moreover,

$$D^* \left( \frac{\partial^3 S}{\partial x^l \partial x^k \partial y^j} \right) (\mathbf{q}) = -\Gamma_{jkl} + \alpha T_{jkl}, \quad (3.10)$$

$$D^* \left( \frac{\partial^3 S}{\partial y^l \partial y^k \partial x^j} \right) (\mathbf{q}) = -\Gamma_{jkl} - \alpha T_{jkl}. \quad (3.11)$$

Hence, choosing  $\alpha = 1/2$  in the Lagrangian solves the inverse problem, since subtracting the latter formulae from each other is precisely one of the definitions in (2.10). Remember that there are other 6 possible linear combinations of the derivatives that lead to rank three tensors. However, one can check that they generate the same tensor  $T$  up to a sign.

Strictly speaking, in order to prove (3.10) and (3.11) it is not necessary to consider the term  $\ddot{\mathbf{v}}$  in the expressions (3.5) and (3.6). Neither is it necessary

to consider the expressions (3.8) and (3.9) up to order  $\mathcal{O}((\mathbf{y} - \mathbf{x})^3)$ . The contributions of these terms to the third order derivatives of the principal function vanish identically after taking the pull-back to the diagonal. In fact, these terms were not taken into account in [10] but have been considered here since we are willing to study what happens in order to generate higher order tensors. In particular, we are interested in the possibility of generating fourth order tensor fields.

## 4 Higher rank tensors in the general case

In this section we will consider the problem of defining tensors of order four out of a two-point function, or more concretely, out of a potential function, cf. Eq. (2.2). Our first objective will be to find linear combinations of derivatives of the two-point function giving raise to rank four tensors. We will use the following abbreviated notation

$$LRLR := D^* (\mathcal{L}_{X_L} \mathcal{L}_{Y_R} \mathcal{L}_{Z_L} \mathcal{L}_{W_R} S).$$

And similarly for other combinations of Left- and Right-lifts. This is convenient because the only differences between the terms will be the particular combination of Left- and Right-lifts that appear. There are only two possible rank four tensors, namely

$$Q_1(X, Y, Z, W) = (LRLR - RLLR) + (LRRR - RLLL) \\ + (LLLL - RLRR) + (LRLR - RLRL), \quad (4.1)$$

$$Q_2(X, Y, Z, W) = (LLLL - RRRR) + (LLLR - RRRL) \\ + (LLRL - RRLR) + (LLRR - RLLL). \quad (4.2)$$

We can now compute the derivatives of the principal function and see what is the rank four tensor that one gets. In principle, one could also enlarge the Lagrangian to accommodate a term of order  $\mathcal{O}(\mathbf{v}^4)$ . Before computing the tensors  $Q_1$  and  $Q_2$  we will consider the Lagrangian given by

$$\mathfrak{L}(\mathbf{q}, \mathbf{v}) = \frac{1}{2} g_{ij} v^i v^j + \frac{\alpha}{6} T_{ijk} v^i v^j v^k + \frac{1}{24} C_{ijkl} v^i v^j v^k v^l, \quad (4.3)$$

where now  $C$  is some fourth order tensor. Using the same analysis of the previous section one can show that the first order derivatives of the principal function become now:

$$\frac{\partial S}{\partial x^i} = -g_{ij}(\mathbf{x}) \Delta^j - \frac{1}{2} \Gamma_{ijk}(\mathbf{x}) \Delta^j \Delta^k - \frac{1}{6} \Gamma_{ijk}(\mathbf{x}) \Gamma_{ls}^j(\mathbf{x}) \Delta^k \Delta^l \Delta^s \\ - \frac{\alpha}{2} T_{ijk}(\mathbf{x}) \Delta^j \Delta^k - \frac{\alpha}{12} A_{ijkl}(\mathbf{x}) \Delta^j \Delta^k \Delta^l \\ - \frac{1}{24} C_{ijkl}(\mathbf{x}) \Delta^j \Delta^k \Delta^l + \mathcal{O}((\mathbf{y} - \mathbf{x})^4). \quad (4.4)$$

$$\frac{\partial S}{\partial y^i} = g_{ij}(\mathbf{y}) \Delta^j - \frac{1}{2} \Gamma_{ijk}(\mathbf{y}) \Delta^j \Delta^k + \frac{1}{6} \Gamma_{ijk}(\mathbf{y}) \Gamma_{ls}^j(\mathbf{y}) \Delta^k \Delta^l \Delta^s \\ + \frac{\alpha}{2} T_{ijk}(\mathbf{y}) \Delta^j \Delta^k - \frac{\alpha}{12} A_{ijkl}(\mathbf{y}) \Delta^j \Delta^k \Delta^l \\ + \frac{1}{24} C_{ijkl}(\mathbf{y}) \Delta^j \Delta^k \Delta^l + \mathcal{O}((\mathbf{y} - \mathbf{x})^4). \quad (4.5)$$

Now one can compute the result of tensors  $Q_1$  and  $Q_2$  and it turns out that both of them vanish identically. This means that this particular potential function does not generate rank four tensors.

In what follows we will show that if the potential function is analytic on a neighbourhood of the diagonal, it cannot generate order four tensors. First notice that condition (2.2) amounts to say that the diagonal of  $\mathcal{M} \times \mathcal{M}$  is a set of critical points for a potential function  $S$ , and there it attains the value  $S(\mathbf{q}, \mathbf{q})$  for  $\mathbf{q} \in \mathcal{M}$ . A Taylor series expansion around a point  $\mathbf{q}_0$  of the diagonal has necessarily the form

$$\begin{aligned} S(\mathbf{x}, \mathbf{y}) = & S(\mathbf{q}_0, \mathbf{q}_0) + h_{ij}(\mathbf{q}_0)(y^i - x^i)(y^j - x^j) + \\ & + t_{ijk}(\mathbf{q}_0)_{ijk}(y^i - x^i)(y^j - x^j)(y^k - x^k) + \\ & + c_{ijkl}(\mathbf{q}_0)(y^i - x^i)(y^j - x^j)(y^k - x^k)(y^l - x^l) + \mathcal{O}((\mathbf{y} - \mathbf{x})^5), \end{aligned} \quad (4.6)$$

where  $\{x^1, x^2, \dots, x^n, y^1, y^2, \dots, y^n\}$  are now coordinates relative to the point  $(\mathbf{q}_0, \mathbf{q}_0)$  on  $\mathcal{M} \times \mathcal{M}$  and where  $h_{ij}$  depends on the second derivatives of  $S$ ,  $t_{ijk}$  depends on the third derivatives of  $S$  and so on. This is so because all the derivatives along the direction tangent to the diagonal vanish.

If one computes the tensors  $Q_1$  and  $Q_2$  in this particular case one gets that they vanish identically. This means that a potential function, assuming it is analytic, is not capable of generating non-trivial covariant tensor fields of order four.

## 5 Conclusions

We have studied the inverse problem for potential functions and found a solution in terms of Hamilton's principal function and a suitable chosen Lagrangian built from the metric and the skewness tensor. The solution will exist provided that the equations of motion of the associate Hamilton-Jacobi problem are completely integrable. We have also studied if one is able to recover higher order rank tensors (order four) from a potential function satisfying (2.2). It turns out that the answer is negative assuming that the potential function is analytic on a neighbourhood of the diagonal.

Notice that potential functions, as they are described here, have the limitation that they cannot generate non-symmetric tensors. For instance, one cannot generate a symplectic structure using the procedure described. One way of achieving that could be by considering Grassmann variables.

Another interesting line to pursue would be that of finding a way of generating higher order tensors. From what we have seen, at least for order four, this will not be achievable by a two-point function. By that we mean that the tensor arises as linear combinations of the derivatives of the potential function. It is clear that once one has obtained the Christoffel symbols, for instance, one can compute out of them the Riemann tensor in the usual way. What we have shown is that the Riemann tensor cannot be directly obtained by computing fourth order derivatives of the potential function. It is very likely that one could do that by considering three-point functions. However, developing such a theory in depth would also require a Hamilton-Jacobi theory of order higher than two (associated to evolution equations including up to third order derivatives).

## Acknowledgements

The authors want to thank the scientific committee of the “XXVI International Fall Workshop on Geometry and Physics” for giving them the opportunity to present this research. G.M. acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Programme for Centres of Excellence in RD (SEV-2015/0554), and would like to thank the support provided by the Santander/UC3M Excellence Chair Programme 2016/2017. J.M.P.P. was supported by QUITEMAD+ S2013/ICE-2801, the Spanish MINECO grant MTM2017-84098-P and the “Juan de la Cierva - Incorporación” Project 2018/00002/001.

## References

- [1] *Differential geometry in statistical inference*. S.I. Amari et. al. Institute of Mathematical Statistics, Hayward, California, 1987.
- [2] *Methods of Information Geometry*. S.I. Amari and H. Nagaoka. American Mathematical Society, Providence, Rhode Island, 2000.
- [3] A. Ashtekar and T.A. Schilling. Geometrical Formulation of Quantum Mechanics. In: Harvey A. (eds) *On Einstein's Path*. Springer, New York, NY, 1999.
- [4] N. Ay and S. I. Amari. A novel approach to canonical divergences within information geometry. *Entropy*, **17** (12), 8111–8129, (2015).
- [5] V. Cantoni. Generalized “Transition Probability”. *Communications in Mathematical Physics*, **44**, 125–128, (1975).
- [6] V. Cantoni. The Riemannian Structure on the States of Quantum-like Systems. *Communications in Mathematical Physics*, **56**, 189–193, (1977).
- [7] F.M. Ciaglia, F. Di Cosmo, G. Marmo. Hamilton-Jacobi Theory and Information Geometry. In: Nielsen F., Barbaresco F. (eds) *Geometric Science of Information*. GSI 2017. Lecture Notes in Computer Science, **10589**. Springer, Cham
- [8] *Statistical Decision Rules and Optimal Inference*. N.N. Cencov. American Mathematical Society, Providence, Rhode Island, 1982.
- [9] F.M. Ciaglia et. al. Dynamical Vector Fields on the Manifold of Quantum States. *Open Syst. Inf. Dyn.* **24**, 1740003, (2017).
- [10] F.M. Ciaglia et. al. Hamilton-Jacobi approach to potential functions in information geometry. *Journal of Mathematical Physics* **58**, 063506, (2017).
- [11] F.M. Ciaglia et. al. A pedagogical intrinsic approach to Relative Entropies as Potential Functions of Quantum metrics: The  $q$ - $z$  family. *ArXiv:1711.09769*.
- [12] I. Csiszár. Information type measures of difference of probability distributions and direct observations. *Studia Scientiarum Mathematicarum Hungarica* **2**, 299–318, (1967).
- [13] ———. On topological properties of  $f$ -divergence. *Studia Scientiarum Mathematicarum Hungarica* **2**, 329–339, (1967).
- [14] J.J. Duistermaat. On Hessian Riemannian Structures. *Asian J. Math.* **5** (1), 79-92, (2001).
- [15] E. Ercolessi, G. Marmo and G. Morandi. From the equations of motion to the canonical commutation relations. *La rivista del Nuovo Cimento della Societ'a Italiana di Fisica*, **33**(8-9), (2010).
- [16] E. Ercolessi and M. Schiavina. Symmetric logarithmic derivative for general n-level systems and the quantum Fisher information tensor for three-level systems. *Phys. Lett. A* **377**, 1996–2002, (2013).
- [17] P. Facchi et. al. Classical and quantum Fisher information in the geometrical formulation of quantum mechanics. *Phys. Lett. A*, **374** (48), 4801–4803, (2010).

- [18] R.A. Fisher. Theory of statistical estimation. *Proc. Cambridge Philos.* **122**, 700–725 (1925).
- [19] A. Fujiwara and H. Nagaoka. Quantum Fisher metric and estimation for pure state models. *Phys. Lett. A* **201**, 119–124, (1995).
- [20] J. Grabowski, M. Kuś and G. Marmo. Geometry of Quantum Systems: density states and entanglement. *J. Phys. A* **38** (47), 10217–10244, (2005).
- [21] C.W. Helstrom. *Quantum Detection and Estimation Theory*. Academic Press, New York, 1976.
- [22] S. Kullback and R.A. Leibler. On information sufficiency. *Annals of Mathematical Statistics* **22** (1), 79–86, (1951).
- [23] C. Lanczos. *The Variational Principles of Mechanics*. Dover Publications, New York (1970).
- [24] V.I. Man’ko, G. Marmo, F. Zaccaria and ECG. Sudarshan. Differential geometry of density states. *Reports on Mathematical Physics* **55** (3), 405–422, (2005).
- [25] G. Morandi, et. al. The inverse problem in the calculus of variations and the geometry of the tangent bundle. *Physics Reports* **188**, (3 & 4), 147–284, (1990).
- [26] D. Petz. Monotone metrics on matrix spaces. *Linear Alg. and its Applications* **244**, 81–96, (1996)
- [27] C.R. Rao. Information and accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.* **37**, 81–91 (1945).
- [28] H. Shima. *The Geometry of Hessian Structures*. World Scientific, New Jersey, 2007.
- [29] *Theory of statistical inference and information*. I. Vajda. Kluwer Academic Publishers, 1989.
- [30] W.K. Wootters. Statistical distance and Hilbert space. *Phys. Rev. D*, **23** (2), 357, (1981).