Chin. Ann. Math.
33B(6), 2012, 889–902
DOI: 10.1007/s11401-012-0745-9

Chinese Annals of Mathematics, Series B

Statistical Structures on Metric Path Spaces*

Mircea CRASMAREANU[1] Cristina-Elena HREŢCANU[2]

Abstract The authors extend the notion of statistical structure from Riemannian geometry to the general framework of path spaces endowed with a nonlinear connection and a generalized metric. Two particular cases of statistical data are defined. The existence and uniqueness of a nonlinear connection corresponding to these classes is proved. Two Koszul tensors are introduced in accordance with the Riemannian approach. As applications, the authors treat the Finslerian (α, β)-metrics and the Beil metrics used in relativity and field theories while the support Riemannian metric is the Fisher-Rao metric of a statistical model.

Keywords Semispray, Nonlinear connection, Metric path space, Statistical structure, Skewness, Koszul tensors, (α, β)-metric, Beil metric, Rayleigh statistical structure, Fisher-Rao metric, Statistical model

2010 MR Subject Classification 53C05, 53C15, 53C60, 53B15, 53B20, 53C80

1 Introduction

Information geometry arose from investigating the geometrical structure of a class of probability distributions depending on various parameters, and was applied successfully to various areas including statistical inference, control system theory and multi-terminal information theory (see [1]). A main notion of this theory is that of the statistical manifold, which has two (equivalent) variants. The first one is a triple (M, g, ∇) with g (a Riemannian metric on the manifold M) and ∇ (a symmetric linear connection for which $C := \nabla g$ is totally symmetric), where C is called the cubic form and the difference between the Levi-Civita connection of g, and ∇ is characterized by a Koszul form (see [13, p. 149]). The second one (see [8]) is a triple (M, g, D) with D, a completely symmetric tensor field of $(0, 3)$-type called skewness. In the present work, we consider the second point of view and recall some of its tools. More precisely, for the pair (g, D), we associate the tensor field $\widetilde{D}$ of $(1, 2)$-type given by

$$g(\widetilde{D}(X, Y), Z) = D(X, Y, Z) \tag{1.1}$$

with the linear connection

$$\overset{\alpha}{\nabla} = \overline{\nabla} - \frac{\alpha}{2}\widetilde{D} \tag{1.2}$$

Manuscript received July 20, 2011. Revised July 18, 2012.
[1]Faculty of Mathematics, Alexandru Ioan Cuza University, Iaşi 700506, Romania.
E-mail: mcrasm@uaic.ro
[2]Faculty of Food Engineering, Stefan Cel Mare University, Suceava 720229, Romania.
E-mail: criselenab@yahoo.com
*Project supported by the Romanian National Authority for Scientific Research, CNCS UEFISCDI (No. PN-II-ID-PCE-2012-4-0131).

for a real number α, where $\overline{\nabla}$ is the Levi-Civita connection of g, and X, Y, Z are vector fields on M. There are three important properties of this setting:

(1) $\overset{\alpha}{\nabla}$ is a torsion-free linear connection (see [1, p. 33]).

(2) $(\overset{\alpha}{\nabla}_X g)(Y, Z) = \alpha D(X, Y, Z)$.

(3) $\overset{\alpha}{\nabla}$ and $\overset{-\alpha}{\nabla}$ are g-conjugated (see [1, p. 52])

$$X(g(Y, Z)) = g(\overset{\alpha}{\nabla}_X Y, Z) + g(X, \overset{-\alpha}{\nabla}_X Z). \tag{1.3}$$

In information geometry, statistical models have a Fisher metric as the Riemannian metric, and the $\overset{1}{\nabla}$ is said to be an exponential connection or e-connection for short, while $\overset{-1}{\nabla}$ is called a mixture connection or m-connection. The statistical model of the exponential family is 1-flat (see [1, p. 35]).

The aim of the present paper is to extend this notion of statistical structure to the geometry of systems of the second order differential equations on M. More precisely, given such a system S, on short semispray, we can obtain a type of differential ∇, if S is considered as a vector field on the tangent bundle TM. A main tool in the definition of ∇ is given by a splitting of the iterated tangent bundle $T(TM)$ provided by a distribution N on TM. Such an object N is called a nonlinear connection. A remarkable result is that every S yields such a nonlinear connection $\overset{c}{N}$ indexed by us with c from canonical one.

We consider a triple (semispray S, nonlinear connection N, generalized metric g) on TM a pair (symmetric two-covariant tensor field D, α), and consider the expectation of D as the skewness of the statistical data. Then we derive the α-version of ∇ in this framework and study its properties, e.g., the variant of (1.3) holds if and only if N is a metric nonlinear connection. Also, we introduce two types of statistical structures, and prove that there exists a unique nonlinear connection, called special, for these cases. A Koszul difference of ∇'s is introduced and it contracts with the metric. We consider a Koszul function, relating the canonical nonlinear connection with the nonlinear connection, belonging to the given statistical structure.

A setting, where we have such pairs (g, D), is the Finslerian geometry of (α, β)-metrics. Let us remark that a relationship between Riemann-Finsler geometry and information geometry was already provided in [12]. More precisely, starting with a Riemannian metric a and a 1-form b both on M, we get on the tangent bundle TM two Riemannian metrics: g_a the Sasaki lift of a as well as the Finsler metric D generated by a and b in the form of (α, β)-metrics. In consequence, we express the special nonlinear connection N of this framework called Finsler statistical data. For example, we treat the case of Randers, Kropina and Riemann-type (α, β)-metrics. The second class of examples consists in Beil metrics, a class of generalized metrics widely used in some physical theories (see [3–4, 10]). The Riemannian metric, on which all these generalized metrics are built, is the Fisher-Rao metric of a statistical model. The case of Gaussian distributions is discussed. Another possible underlying Riemannian metric is the one considered by Shen in [12] as generated by an f-divergence.

Section 5 adds a new covariant tensor field inspired by the control theory, which is called in correspondence with this domain Rayleigh dissipation. We obtain the entire family of nonlinear connections adapted via the α-dynamical derivative to this new structure. In particular, we derive the nonlinear connections making α-parallel the given generalized metric g.

2 Nonlinear Connections and Semisprays on Tangent Bundles

Let M be a smooth, n-dimensional manifold, for which we denote by $C^\infty(M)$ the algebra of smooth real functions on M; by $\mathcal{X}(M)$ the Lie algebra of vector fields on M; by $T^r_s(M)$ the $C^\infty(M)$-module of tensor fields of (r,s)-type on M.

A local chart $x = (x^i) = (x^1, \cdots, x^n)$ on M lifts to a local chart on the tangent bundle TM given by $(x,y) = (x^i, y^i)$. If $\pi : TM \to M$ is the canonical projection, then the kernel of the differential of π is an integrable distribution $V(TM)$ with a local basis $(\frac{\partial}{\partial y^i})$. An important element of $V(TM)$ is the Liouville vector field $\mathbb{C} = y^i \frac{\partial}{\partial y^i}$. $V(TM)$ is called the vertical distribution and its elements are vertical vector fields.

The tensor field $J \in T^1_1(TM)$ given by $J = \frac{\partial}{\partial y^i} \otimes \mathrm{d}x^i$ is called the tangent structure. Two of its properties are the nilpotence $J^2 = 0$ and $\operatorname{im} J (= \ker J) = V(TM)$, respectively.

A well-known notion in the tangent bundles geometry is as follows.

Definition 2.1 (see [6]) *A supplementary distribution N to the vertical distribution $V(TM)$*

$$T(TM) = N \oplus V(TM) \tag{2.1}$$

is called a horizontal distribution or a nonlinear connection. A vector field belonging to N is called to be horizontal.

A nonlinear connection has a local basis

$$\frac{\delta}{\delta x^i} := \frac{\partial}{\partial x^i} - N^j_i \frac{\partial}{\partial y^j}, \tag{2.2}$$

and the functions $(N^i_j(x,y))$ are called the coefficients of N. So, a basis of $\mathcal{X}(TM)$ adapted to the decomposition (2.1) is $(\frac{\delta}{\delta x^i}, \frac{\partial}{\partial y^i})$ called the Berwald basis. The dual of the Berwald basis is $(\mathrm{d}x^i, \delta y^i = \mathrm{d}y^i + N^i_j \mathrm{d}x^j)$.

The second remarkable structure on TM is provided by the following definition.

Definition 2.2 (see [6]) $S \in \mathcal{X}(TM)$ *is called a semispray, if*

$$J(S) = \mathbb{C}. \tag{2.3}$$

In the canonical coordinates, the semispray S has the form

$$S = y^i \frac{\partial}{\partial x^i} - 2G^i(x,y) \frac{\partial}{\partial y^i}, \tag{2.4}$$

where the functions $(G^i(x,y))$ are the coefficients of S. The flow of S is a system of second order differential equations $\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} = 2G^i(x, \frac{\mathrm{d}x}{\mathrm{d}t})$. Then the pair (M,S) is called a path space (see [11]).

An important remark is that a nonlinear connection $N = (N^i_j)$ yields a unique horizontal semispray denoted $S(N)$ with

$$G^i = \frac{1}{2} N^i_j y^j, \tag{2.5}$$

or in other words,

$$S(N) = y^i \frac{\delta}{\delta x^i}. \tag{2.6}$$

Conversely, a semispray S yields a nonlinear connection $\overset{c}{N}$ given by

$$\overset{c}{N}{}^i_j = \frac{\partial G^i}{\partial y^j}. \tag{2.7}$$

Definition 2.3 *A semispray S, for which the coefficients (G^i) are homogeneous and are of degree 2 with respect to the variables (y^i), is called a spray.*

Locally, this means that, via the Euler theorem, we have

$$2G^i = y^j \frac{\partial G^i}{\partial y^j}. \tag{2.8}$$

Then $\overset{c}{N}$ is 1-homogeneous, and

$$\overset{c}{N}{}^i_j = y^a \frac{\partial \overset{c}{N}{}^i_j}{\partial y^a}, \tag{2.9}$$

which yields that S is horizontal with respect to $\overset{c}{N}$, i.e., S has the expression (2.6).

The weak torsion of the nonlinear connection N is the vertical valued 2-form

$$t(X,Y) = J[hX, hY] - v[hX, JY] - v[JX, hY], \tag{2.10}$$

or in local coordinates,

$$t = \frac{1}{2}\Big(\frac{\partial N^i_j}{\partial y^k} - \frac{\partial N^i_k}{\partial y^j}\Big) \mathrm{d}x^k \wedge \mathrm{d}x^j \otimes \frac{\partial}{\partial y^i}. \tag{2.11}$$

The nonlinear connection N is called to be symmetric if $t = 0$.

3 Statistical Structures for Metric Path Spaces

Let us fix a semispray $S = (G^i)$ and a nonlinear connection $N = (N^i_j)$. Following [6], let us consider the following definition.

Definition 3.1 *The dynamical derivative associated with the pair (S, N) is the map $\overset{SN}{\nabla}$: $N \to N$ given by*

$$\overset{SN}{\nabla} X = \overset{SN}{\nabla} \left(X^i \frac{\delta}{\delta x^i}\right) := (S(X^i) + N^i_j X^j) \frac{\delta}{\delta x^i}. \tag{3.1}$$

The dynamical derivative associated with $(S, \overset{c}{N})$ is denoted by $\overset{S}{\nabla}$.

Some properties of this geometrical object are as follows:

(I) $\overset{SN}{\nabla} \left(\frac{\delta}{\delta x^i}\right) = N^j_i \frac{\delta}{\delta x^j}$.

(II) $\overset{SN}{\nabla} (X + Y) = \overset{SN}{\nabla} X + \overset{SN}{\nabla} Y$.

(III) $\overset{SN}{\nabla} (fX) = S(f)X + f \overset{SN}{\nabla} X$.

It is straightforward to extend the action of $\overset{SN}{\nabla}$ to general horizontal tensor fields by the preservation of tensor products and the Leibniz rule. Moreover, we will extend the $\overset{SN}{\nabla}$ to a special class of tensor fields.

Definition 3.2 *A d-tensor field (d from distinguished) on TM is a tensor field, whose changes of components, under a change of canonical coordinates $(x,y) \to (\widetilde{x},\widetilde{y})$ on TM, involve only factors of type $\frac{\partial \widetilde{x}}{\partial x}$ and/or $\frac{\partial x}{\partial \widetilde{x}}$.*

Example 3.1 (i) $\left(\frac{\delta}{\delta x^i}\right)$ and $\left(\frac{\partial}{\partial y^i}\right)$ are components of d-tensor fields of $(1,0)$-type.

(ii) $(\mathrm{d}x^i)$ and (δy^i) are components of d-tensor fields of $(0,1)$-type.

(iii) (G^i) are not components of a d-tensor field, since a change of coordinates implies

$$2\widetilde{G}^i = 2\frac{\partial \widetilde{x}^i}{\partial x^j} G^j - \frac{\partial \widetilde{y}^i}{\partial x^j} y^j.$$

But the result is that given two semisprays $\overset{1}{S}$ and $\overset{2}{S}$, their difference $X = \overset{2}{S} - \overset{1}{S}$ is a vertical vector field (and then a vertical d-vector field).

(iv) (N^i_j) are not components of a d-tensor field, since a change of coordinates implies

$$\frac{\partial \widetilde{x}^j}{\partial x^k} N^k_i = \widetilde{N}^j_k \frac{\partial \widetilde{x}^k}{\partial x^i} + \frac{\partial \widetilde{y}^j}{\partial x^i}.$$

It follows that, given two nonlinear connections $\overset{1}{N}$ and $\overset{2}{N}$, their difference $F = \overset{2}{N} - \overset{1}{N} = (F^i_j = \overset{2}{N}{}^i_j - \overset{1}{N}{}^i_j)$ is a d-tensor field of $(1,1)$-type. Therefore, the set $\mathcal{N}(S,g)$ of all nonlinear connections is a $C^\infty(TM)$-affine module associated with the $C^\infty(TM)$-linear module of d-tensor fields of $(1,1)$-type.

Definition 3.3 *A (generalized) metric g on TM is a d-tensor field of $(0,2)$-type, which is symmetric and non-degenerated. The datum (M,S,N,g) is called an N-metric path space. In particular, the $\overset{c}{N}$-metric path space is called a metric path space.*

For the components $g_{ij} = g\left(\frac{\delta}{\delta x^i}, \frac{\delta}{\delta x^j}\right)$, the following properties hold:

(1) (Symmetry) $g_{ij} = g_{ji}$.

(2) (Non-degeneration) $\det(g_{ij}) \neq 0$, then there exists a d-tensor field of $(2,0)$-type $g^{-1} = (g^{ij})$.

The definition is justified from the fact that $g_{ij}\mathrm{d}x^i \otimes dx^j + g_{ij}\delta y^i \otimes \delta y^j$ is a Riemannian metric on TM for which N and $V(TM)$ are orthogonal distributions.

Definition 3.4 *The dynamical derivative of the metric g associated with the pair (S,N) is the map $\overset{SN}{\nabla} g : N \times N \to N$ given by*

$$\overset{SN}{\nabla} g(X,Y) = S(g(X,Y)) - g(\overset{SN}{\nabla} X, Y) - g(X, \overset{SN}{\nabla} Y). \tag{3.2}$$

According with [5], *the nonlinear connection is called a metric nonlinear connection if $\overset{SN}{\nabla} g = 0$.*

The main notion of this section is the following definition.

Definition 3.5 *A statistical structure on the N-metrical space is a pair (a symmetric d-tensor field D of $(0,2)$-type called skewness, a real number α). We consider also the d-tensor field $\widetilde{D}$ of $(1,1)$-type determined by g and D (the skewness operator)*

$$g(\widetilde{D}(X), Y) = D(X,Y). \tag{3.3}$$

For a statistical structure, we consider the map $\overset{\alpha}{\nabla}: N \to N$ given by

$$\overset{\alpha}{\nabla} = \overset{SN}{\nabla} - \frac{\alpha}{2}\widetilde{D}, \tag{3.4}$$

which we call the α-dynamical derivative of the statistical datum (M, S, N, g, D, α). Let $D_{ij} = D\left(\frac{\delta}{\delta x^i}, \frac{\delta}{\delta x^j}\right)$ be the local components of the skewness tensor field, and $\widetilde{D}^i_j = g^{ia}D_{aj}$ be its $(1,1)$-type variant (2.3). The properties of the α-dynamical derivative are given by the following theorem.

Theorem 3.1 *The α-dynamical derivative satisfies:*
(α-I) $\overset{\alpha}{\nabla}\left(\frac{\delta}{\delta x^i}\right) = \left(N^i_j - \frac{\alpha}{2}g^{ja}D_{ai}\right)\frac{\delta}{\delta x^j}$.
(α-II) $\overset{\alpha}{\nabla}(X+Y) = \overset{\alpha}{\nabla} X + \overset{\alpha}{\nabla} Y$.
(α-III) $\overset{\alpha}{\nabla}(fX) = S(f)X + f\overset{\alpha}{\nabla} X$.
The α-dynamical derivative of the metric is

$$\overset{\alpha}{\nabla} g = \overset{SN}{\nabla} g + \alpha D. \tag{3.5}$$

Proof We only prove (3.5). Similar to (3.2), we have

$$\begin{aligned}(\overset{\alpha}{\nabla} g)(X,Y) &= S(g(X,Y)) - g(\overset{\alpha}{\nabla} X, Y) - g(X, \overset{\alpha}{\nabla} Y)\\ &= (\overset{SN}{\nabla} g)(X,Y) + \frac{\alpha}{2}(D(X,Y) + D(Y,X)),\end{aligned}$$

which gives us the conclusion.

Let us remark that (3.5) is our version of property (2) in Section 1 (which is completely recovered if N is a metric nonlinear connection), while the relation (1.3) becomes that as in the following theorem.

Theorem 3.2 *The statistical data (M, S, N, g, D, α) and $(M, S, N, g, D, -\alpha)$ are in duality with respect to S,*

$$S(g(X,Y)) = g(\overset{\alpha}{\nabla} X, Y) + g(X, \overset{-\alpha}{\nabla} Y) \tag{3.6}$$

if and only if N is a metric nonlinear connection.

Proof We have

$$\begin{aligned}g(\overset{\alpha}{\nabla} X, Y) + g(X, \overset{-\alpha}{\nabla} Y) &= g(\overset{SN}{\nabla} X, Y) + g(X, \overset{SN}{\nabla} Y) - \frac{\alpha}{2}(g(\widetilde{D}(X), Y) - g(X, \widetilde{D}(Y))\\ &= S(g(X,Y)) - \overset{SN}{\nabla} g(X,Y) - \frac{\alpha}{2}(D(X,Y) - D(Y,X)),\end{aligned}$$

which yields the conclusion.

We introduce two particular cases of the general framework presented above.

Definition 3.6 *The statistical datum (M, S, N, g, D, α) is called*

(i) *self-dual, if the skewness D is a conformal deformation of g, which means that there exists a smooth function $\rho \in C^\infty(TM)$, such that $D = \rho g$,*

(ii) *(β, γ)-special with $\beta, \gamma \in \mathbb{R}$, if*

$$\overset{SN}{\nabla} = \beta\widetilde{D} + \gamma I. \tag{3.7}$$

It follows that for a self-dual statistical datum, we have that $\widetilde{D} = \rho I$ with $I = (\delta^i_j)$, i.e., the Kronecker delta tensor, and the α-dynamical derivative of the metric is $\overset{\alpha}{\nabla} g = \overset{SN}{\nabla} g + \alpha\rho g$.

A (β, γ)-special statistical datum has the α-dynamical derivative on the metric

$$\overset{\alpha}{\nabla} g = S(g(\,\cdot\,,\,\cdot\,)) - (2\beta - \alpha)D - 2\gamma g. \tag{3.8}$$

A (β, γ)-special self-dual statistical datum has the α-dynamical derivative

$$\overset{\alpha}{\nabla} = \Big[\Big(\beta - \frac{\alpha}{2}\Big)\rho + \gamma\Big] I. \tag{3.9}$$

In particular, $\overset{2\beta}{\nabla} = \gamma I$ and a direct computation yield the following theorem.

Theorem 3.3 *Given a statistical datum $(M, S, g, D, \beta, \gamma)$, there exists a unique nonlinear connection (denoted by $\overset{s}{N}$), such that, the (β, γ)-special is the given datum. Its coefficients are*

$$\overset{s}{N}{}^i_j = \beta g^{ia} D_{aj} + \gamma \delta^i_j. \tag{3.10}$$

If $\beta \neq 0$, then $\overset{s}{N}$ is symmetric if and only if

$$\frac{\partial(g^{ia} D_{aj})}{\partial y^k} = \frac{\partial(g^{ia} D_{ak})}{\partial y^j}, \tag{3.11}$$

while for $\beta = 0$, we have that $\overset{s}{N}$ is symmetric. In addition, if $(M, S, g, D, \beta, \gamma)$ is self-dual, then we denote the above nonlinear connection by $\overset{sd}{N}$ with

$$\overset{sd}{N}{}^i_j = (\beta\rho + \gamma)\delta^i_j. \tag{3.12}$$

If $\beta \neq 0$, then $\overset{sd}{N}$ is symmetric if and only if ρ is a constant.

Let us recall that a special is the nonlinear connection (3.10) and a special dual is the nonlinear connection (3.11). Both these special nonlinear connections satisfy a symmetry with respect to the metric g,

$$g_{iu} \overset{\cdot}{N}{}^u_j = g_{ju} \overset{\cdot}{N}{}^u_i . \tag{3.13}$$

Theorem 3.4 *$\overset{s}{N}$ is a metric nonlinear connection if and only if*

$$S(g_{ij}) = 2(\beta D_{ij} + \gamma g_{ij}), \tag{3.14}$$

while $\overset{sd}{N}$ is a metric nonlinear connection if and only if

$$S(g_{ij}) = 2(\beta\rho + \gamma) g_{ij}. \tag{3.15}$$

Let us end this section with a Koszul type approach. More precisely, for our framework, we give the following definition.

Definition 3.7 *The Koszul tensor of the statistical datum* (M, S, N, g, D, α) *is* $K_\alpha \in T^0_2(TM)$ *given by*

$$K_\alpha = \overset{\alpha}{\nabla} g - \overset{S}{\nabla} g. \tag{3.16}$$

The Koszul function $k_\alpha \in C^\infty(TM)$ *is*

$$k_\alpha = \mathrm{Trace}_g(K_\alpha). \tag{3.17}$$

By

$$K_\alpha = (\overset{SN}{\nabla} - \overset{S}{\nabla})g + \alpha D, \tag{3.18}$$

we derive

$$k_\alpha = \sum_{i=1}^{n}[2(N^i_i - \overset{c}{N}{}^i_i) + \alpha g^{ia} D_{ai}]. \tag{3.19}$$

Therefore, in the self-dual case,

$$k_\alpha = 2\sum_{i=1}^{n}[\overset{sd}{N}{}^i_i - \overset{c}{N}{}^i_i] + n\alpha\rho, \tag{3.20}$$

in the particular case of self-dual (β, γ)-special datum,

$$k_\alpha = 2\Big[n(\beta\rho + \gamma) - \sum_{i=1}^{n} \overset{c}{N}{}^i_i\Big] + n\alpha\rho. \tag{3.21}$$

4 Examples

4.1 Finsler (α, β)-metrics

Let us recall the following definition.

Definition 4.1 *A Finsler fundamental function is a map* $F : TM \to \mathbb{R}_+$, *such that*

(F1) F *is smooth on the slit tangent bundle* $T_0M = TM \setminus \{0\}$ *and is continuous on the null section* $\{0\}$ *of the projection* $\pi : TM \to M$.

(F2) F *is positive homogeneous of order one with respect to the fibre coordinates, i.e.,* $F(x, \lambda y) = \lambda F(x, y)$ *for* $\lambda > 0$.

(F3) *For any* $(x, y) \in T_0M$, *the symmetric bilinear form* $D(x, y)$, *satisfying*

$$D_{(x,y)}(u, v) = \frac{1}{2}\frac{\partial^2}{\partial s \partial t}[F^2(x, y + su + tv)]\big|_{s=t=0}, \quad y, u, v \in T_xM, \tag{4.1}$$

is non-degenerated and has constant signature.

By homogeneity, it holds that $F^2 = D_{ij}y^i y^j$ with

$$D_{ij} = \frac{1}{2}\frac{\partial^2 F^2}{\partial y^i \partial y^j}. \tag{4.2}$$

Then $D = (D_{ij})$ is called the Finsler metric generated by F.

We consider D the Christoffel symbols:

$$\gamma^i_{jk}(x,y)=\frac{1}{2}D^{ia}\Big(\frac{\partial D_{ak}}{\partial x^j}+\frac{\partial D_{ja}}{\partial x^k}-\frac{\partial D_{jk}}{\partial x^a}\Big). \tag{4.3}$$

Then we obtain the Finslerian spray S_F with

$$G^i=\frac{1}{2}\gamma^i_{jk}y^jy^k. \tag{4.4}$$

The canonical nonlinear connection is called the Cartan nonlinear connection, which is a metrical nonlinear connection with the expression

$$\overset{c}{N}{}^i_j=\gamma^i_{jk}y^k-C^i_{mj}\gamma^m_{ks}y^ky^s \tag{4.5}$$

and the vertical Christoffel symbols $C^i_{jk}=D^{ia}C_{ajk}$, where

$$C_{ajk}=\frac{1}{4}\frac{\partial^3F^2}{\partial y^a\partial y^j\partial y^k}=\frac{1}{2}\frac{\partial D_{jk}}{\partial y^a}. \tag{4.6}$$

Consider now a Riemannian metric $a=(a_{ij}(x))$ on the base manifold M and $b=(b_i(x))$ of 1-form also on M, with which we associate the function $\alpha(x,y)=\sqrt{a_{ij}y^iy^j}$ and the function $\beta(x,y)=b_iy^i$.

Definition 4.2 *The Finsler space (M,F) is of (α,β)-type, if there exists a 2-homogeneous function $L(\,\cdot\,,\,\cdot\,)$ of two variables, such that*

$$F^2=L(\alpha,\beta). \tag{4.7}$$

By the Riemannian metric a to TM, we obtain the Riemann-Sasaki metric

$$g_a=a_{ij}\mathrm{d}x^i\otimes\mathrm{d}x^j+a_{ij}\mathrm{d}y^i\otimes\mathrm{d}y^i\otimes\mathrm{d}y^j, \tag{4.8}$$

which will be considered in pair with the Finsler-Sasaki metric given in (4.2),

$$D=D_{ij}\mathrm{d}x^i\otimes\mathrm{d}x^j+D_{ij}\mathrm{d}y^i\otimes\mathrm{d}y^j. \tag{4.9}$$

For a Finsler space of (α,β)-type, we consider the following four invariants (see [9, p. 890]):

$$\begin{cases} p=\dfrac{1}{2\alpha}\dfrac{\partial F^2}{\partial\alpha}, & p_0=\dfrac{1}{2}\dfrac{\partial^2F^2}{\partial\beta^2},\\ p_1=\dfrac{1}{2\alpha}\dfrac{\partial^2F^2}{\partial\alpha\partial\beta}, & p_2=\dfrac{1}{2\alpha^2}\Big(\dfrac{\partial^2F^2}{\partial\alpha^2}-\dfrac{1}{\alpha}\dfrac{\partial F^2}{\partial\alpha}\Big),\end{cases} \tag{4.10}$$

where the subscripts denote the minus of degree of homogeneity of these invariants, which connect the Riemannian metric a with the Finsler metric D through the following relation:

$$D_{ij}=pa_{ij}+p_0b_ib_j+p_1(b_iy_j+b_jy_i)+p_2y_iy_j, \tag{4.11}$$

where $y_i=a_{ij}(x)y^j$.

Applying Theorem 3.3, we get the following theorem.

Theorem 4.1 *The special nonlinear connection* $\overset{s}{N}$ *of the Finsler statistical datum* (M, S_F, g_a, D) *is given by*

$$\overset{s}{N}{}^i_j = (\beta p + \gamma)\delta^i_j + \beta[p_0 b^i b_j + p_1(b^i y_j + b_j y^i) + p_2 y^i y_j]. \tag{4.12s}$$

For $\beta \cdot \gamma \neq 0$*, the nonlinear connection* $\overset{s}{N}$ *is a metric nonlinear connection if and only if* $a = (a_{ij})$ *is a constant Riemannian metric and* $b = 0$*. Then*

$$\overset{s}{N}{}^i_j = (\beta p + \gamma)\delta^i_j + \beta p_2 y^i y_j, \tag{4.12sm}$$

and this $\overset{s}{N}$ *is symmetric if and only if* p *and* p_2 *are constants. Moreover,* $\overset{s}{N}$ *is special dual with* $\rho = \frac{-\gamma}{\beta}$*.*

Proof The first part is a direct consequence of (3.10) and (4.11). For the second part, by the condition (3.14), we have

$$y^u \frac{\partial a_{ij}}{\partial x^u} = 2(\beta D_{ij} + \gamma a_{ij}),$$

where the left-hand side is 1-homogeneous in y, while the right-hand side is 0-homogenous in y. Therefore, a does not depend on x, and D is in fact homotetic with a, i.e., $D = \frac{-\gamma}{\beta} a$, which is only possible in Riemannian geometry.

Example 4.1 (1) Randers metrics: $F^2 = (\alpha + \beta)^2$,

$$p = 1 + \frac{\beta}{\alpha}, \quad p_0 = 1, \quad p_1 = \frac{1}{\alpha}, \quad p_2 = -\frac{\beta}{\alpha^3}. \tag{4.13}$$

(2) Kropina: $F^2 = \frac{\alpha^4}{\beta^2}$,

$$p = \frac{2\alpha^2}{\beta^2}, \quad p_0 = \frac{3\alpha^4}{\beta^4}, \quad p_1 = -\frac{4\alpha^2}{\beta^3}, \quad p_2 = \frac{4}{\beta^2}. \tag{4.14}$$

(3) "Riemann" type (α, β)-metric: $F^2 = 1 + \alpha^2$,

$$p = p_0 = 1, \quad p_1 = p_2 = 0. \tag{4.15}$$

4.2 Beil metrics

Consider the metric $g = (g_{ij})$ in the sense of Definition 3.3 and two functions $a, b \in C^\infty(TM)$ with $a \neq 0$ and $b \geq 0$. Let $B = B_i(x, y)\mathrm{d}x^i$ be a vertical 1-form. It holds that

$$D_{ij} = a g_{ij} + b B_i B_j \tag{4.16}$$

is a new metric called the Beil metric or sometimes the Beil deformation of the metric g. The case of semi-Riemannian g (more precisely, Minkowski or Lorentz) on the base M and $a = 1$ with various choices of b and B, was introduced and studied by Beil by constructing a new unified filed theory in [3–4].

Applying Theorem 3.3, we get the following theorem.

Theorem 4.2 *The special dual nonlinear connection* $\overset{sd}{N}$ *of the Beil statistical datum* (M, S, g, D) *is given by*

$$\overset{sd}{N}{}^i_j = (\beta a + \gamma)\delta^i_j + \beta b B^i B_j. \tag{4.17}$$

Example 4.2 (i) The classical Beil metrics: $g = g(x)$.

(ii) Miron-Tavakol metrics (useful in General Relativity): $g = g(x)$, $a = \exp(2\sigma(x, y))$ and $b = 0$. Then we have

$$\overset{s}{N}{}^i_j = (\beta a + \gamma)\delta^i_j. \tag{4.18}$$

4.3 Relationship with statistical models

Let us consider a family M of probability distributions on a set U, such that each element of M can be parametrized by using n real variables $(x^1, \cdots, x^n)$. Then, $M = \{p_x = p(u; x) \mid u \in U, x = (x^1, \cdots, x^n)\}$ and referring to [1, p. 26], we call M an n-dimensional statistical model. So, M is an n-dimensional manifold with the Riemannian metric $a = (a_{ij}(x))$ given by the Fisher information matrix (see [1, p. 28]) or the Fisher-Rao metric

$$a_{ij}(x) = \int p(u, x)\frac{\partial \log p}{\partial x^i}\frac{\partial \log p}{\partial x^j}\mathrm{d}u. \tag{4.19}$$

Example 4.3 Let us consider the family $M = \mathcal{N}(\mu, \sigma^2)$ of 1-dimensional Gaussian probability distributions with mean μ and variance σ^2 on $U = \mathbb{R}$,

$$p(u; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}\exp\Big\{-\frac{(u-\mu)^2}{2\sigma^2}\Big\}. \tag{4.20}$$

Therefore, M is a 2-dimensional manifold parametrized by $\mu \in \mathbb{R}$ and $\sigma \in (0, +\infty)$. We obtain the Fisher-Rao metric (see [14])

$$a = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{pmatrix}. \tag{4.21}$$

Let us remark that the Fisher-Rao metric is different from Shen's metric in [12, p. 92] given by the f-divergence,

$$g = \begin{pmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^2(\sigma^2 + 2\mu^2) \end{pmatrix}. \tag{4.22}$$

Another very interesting geometry is that of gamma distributions studied in details in [2, 12].

5 Rayleigh Statistical Structures

We add to our framework a symmetric d-tensor field of $(0, 2)$-type on TM, denoted by H and called a Rayleigh dissipation due to [7, p. 198], where, in addition, H is considered to be positive-semidefinite. We adopt this general definition without any constraint to the signature of H.

Definition 5.1 *The statistical datum* (M, S, N, g, D, α) *is called an* H*-Rayleigh structure, if the following recurrence relation holds:*

$$\overset{\alpha}{\nabla} g = H. \tag{5.1}$$

The aim of this section is to find all nonlinear connections, which together with fixed (S, g, D, α, H) form a Rayleigh structure. Let us denote by $\mathcal{N}(S, g, D, \alpha, H)$ this family. In order to answer this question, considering Example 3.1(iv), we find that it is necessary to study the following two operators called Obata, acting on the space of d-tensor fields of $(1, 1)$-type:

$$O^{ij}_{kl} = \frac{1}{2}(\delta^i_k\delta^j_l - g^{ij}g_{kl}), \quad \overset{*}{O}{}^{ij}_{kl} = \frac{1}{2}(\delta^i_k\delta^j_l + g^{ij}g_{kl}). \tag{5.2}$$

The Obata operators are supplementary projectors and satisfy

$$O^{ia}_{bj}\overset{*}{O}{}^{bk}_{la} = \overset{*}{O}{}^{ia}_{bj}O^{bk}_{la} = 0, \quad O^{ia}_{bj}O^{bk}_{la} = O^{ik}_{lj}, \quad \overset{*}{O}{}^{ia}_{bj}\overset{*}{O}{}^{bk}_{la} = \overset{*}{O}{}^{ik}_{lj}. \tag{5.3}$$

Then the tensorial equations involving these operators have solutions as indicated in the following theorem.

Theorem 5.1 *The system of equations*

$$\overset{*}{O}{}^{ia}_{bj}(X^b_a) = A^i_j \quad (O^{ia}_{bj}(X^b_a) = A^i_j) \tag{5.4}$$

with unknown X has a solution if and only if

$$O^{ia}_{bj}(A^b_a) = 0 \quad (\overset{*}{O}{}^{ia}_{bj}(A^b_a) = 0). \tag{5.5}$$

Then, the general solution is

$$X^i_j = A^i_j + O^{ia}_{bj}(Y^b_a) \quad (X^i_j = A^i_j + \overset{*}{O}{}^{ia}_{bj}(Y^b_a)) \tag{5.6}$$

with Y, an arbitrary d-tensor field of $(1, 1)$-type.

We are ready for the main result of this section.

Theorem 5.2 *Set S and (g, D, α, H) as above. The family $\mathcal{N}(S, g, D, \alpha, H)$ is given by*

$$N^i_j = \frac{1}{2}\overset{c}{N}{}^i_j - \frac{1}{2}g^{ia}g_{jb}\overset{c}{N}{}^b_a + \frac{1}{2}g^{ia}S(g_{aj}) - \frac{1}{2}g^{ia}(H_{aj} - \alpha D_{aj}) + O^{ia}_{bj}(X^b_a) \tag{5.7}$$

with $X = (X^b_a)$, an arbitrary d-tensor field of $(1, 1)$-type. Therefore, writing

$$N = \overset{c}{N} + \frac{1}{2}[g^{-1}(\overset{S}{\nabla} g - H) + \alpha\widetilde{D}] + O(X), \tag{5.7$'$}$$

we have that $\mathcal{N}(S, g, D, \alpha, H)$ is an affine submodule of $\mathcal{N}(TM)$ passing through the nonlinear connection $\overset{c}{N} + \frac{1}{2}[g^{-1}(\overset{S}{\nabla} g - H) + \alpha\widetilde{D}]$ and having the direction given by the linear submodule $\mathrm{Im}\, O$ of $T^1_1(TM)$.

Proof We search (N^i_j) of the form

$$N^i_j = \overset{c}{N}{}^i_j + F^i_j \tag{5.8}$$

with (F^i_j), a d-tensor field of $(1, 1)$-type to be determined. The local expression of equation (5.1) is

$$S(g_{uv}) - g_{um}N^m_v - g_{mv}N^m_u = H_{uv} - \alpha D_{uv}. \tag{5.9}$$

Inserting (5.8) in (5.9), we have

$$S(g_{uv}) - g_{um} \overset{c}{N}{}^m_v - g_{mv} \overset{c}{N}{}^m_u = g_{um}F^m_v + g_{mv}F^m_u + H_{uv} - \alpha D_{uv}.$$

Multiplying the last relation with g^{ku}, we get

$$g^{ku}S(g_{uv}) - \overset{c}{N}{}^k_v - g^{ku}g_{mv} \overset{c}{N}{}^m_u - g^{ku}(H_{uv} - \alpha D_{uv}) = F^k_v + g^{ku}g_{mv}F^m_u = 2 \overset{*}{O}{}^{kb}_{av} (F^a_b). \quad (5.10)$$

Let us search for the condition similar to (5.5). Then we have

$$\begin{aligned} &O^{kb}_{av}(g^{am}S(g_{mb}) - \overset{c}{N}{}^a_b - g^{am}g_{bl} \overset{c}{N}{}^l_m - g^{am}H_{mb} + \alpha \widetilde{D}^a_b) \\ &= g^{km}S(g_{mv}) - \overset{c}{N}{}^k_v - g^{km}g_{vl} \overset{c}{N}{}^l_m - g^{km}S(g_{mv}) + g^{km}g_{vl} \overset{c}{N}{}^l_m + \overset{c}{N}{}^k_v = 0. \end{aligned}$$

It follows that

$$F^i_j = \frac{1}{2}g^{im}S(g_{mj}) - \frac{1}{2} \overset{c}{N}{}^i_j - \frac{1}{2}g^{ia}g_{jb} \overset{c}{N}{}^b_a - \frac{1}{2}g^{ia}(H_{aj} - \alpha D_{aj}) + O^{ib}_{aj}(X^a_b).$$

Returning to (5.8), we have the conclusion.

In the spray case, equation (5.7) admits a simplification.

Theorem 5.3 *Suppose that S is a spray. The family $\mathcal{N}(S, g, D, \alpha, H)$ is*

$$N^i_j = \frac{1}{2} \overset{c}{N}{}^i_j - \frac{1}{2}g^{ia}g_{jb} \overset{c}{N}{}^b_a + \frac{1}{2}g^{ia}y^m \frac{\delta g_{aj}}{\delta x^m} - \frac{1}{2}g^{ia}(H_{aj} - \alpha D_{aj}) + O^{ia}_{bj}(X^b_a). \quad (5.11)$$

Remark 5.1 (i) The Obata operators split the space of d-tensor fields of $(1,1)$-type into a g-symmetric part $\operatorname{Im} O^* = \operatorname{Ker} O$ with dimension $\frac{n(n-1)}{2}$ and a g-skew-symmetric part $\operatorname{Im} O = \operatorname{Ker} O^*$ with dimension $\frac{n(n+1)}{2}$. The general formula $(5.7)'$ implies that the recurrence relation (5.1) fixes the symmetric part of the tensor field $N - \overset{c}{N}$ as $\frac{1}{2}(\overset{S}{\nabla} g - H + \alpha D)$. An interesting open problem is to consider remarkable geometrical conditions, which fixes the skew-symmetrical part.

(ii) Choosing $H = 0$, we have that, for a given (S, g, D, α), there exists a set of nonlinear connections parametrized by the d-elements of $T^1_1(TM)$, such that g is parallel with respect to $\overset{\alpha}{\nabla}$, i.e., $\overset{\alpha}{\nabla} g = 0$.

References

[1] Amari, S.-I. and Nagaoka, H., Methods of Information Geometry, Translations of Mathematical Monographs, Vol. 191, A. M. S., Providence, RI, 2000.

[2] Arwini, K. A. and Dodson, C. T. J., Information Geometry, Near Randomness and Near Independence, A. J. Doig, W. W. Sampson, J. Scharcanski, et al. (eds.), Lecture Notes in Mathematics, 1953, Springer-Verlag, Berlin, 2008.

[3] Beil, R. G., Electrodynamics from a metric, *Internat. J. Theoret. Phys.*, **26**(2), 1987, 189–197.

[4] Beil, R. G., New class of Finsler metrics, *Internat. J. Theoret. Phys.*, **28**(6), 1989, 659–667.

[5] Bucataru, I., Metric nonlinear connections, *Diff. Geom. Appl.*, **25**(3), 2007, 335–343.

[6] Bucataru, I., Constantinescu, O. and Dahl, M. F., A geometric setting for systems of ordinary differential equations, *Int. J. Geom. Meth. Mod. Phys.*, **8**(6), 2011, 1291–1327.

[7] Bullo, F. and Lewis, A. D., Geometric Control of Mechanical Systems. Modeling, Analysis, and Design for Simple Mechanical Control Systems, Texts in Applied Mathematics, Vol. 49, Springer-Verlag, New York, 2005.

[8] Lauritzen, S. L., Statistical manifolds, Differential Geometry in Statistical Inference, Institute of Mathematical Statistics Lecture Notes Monograph Series, Vol. 10, S.-I. Amari et al. (eds.), Institute of Mathematical Statistics, Hayward, CA, 1987, 163–216.

[9] Matsumoto, M., Finsler geometry in the 20th-century, Handbook of Finsler Geometry, Vol. 1–2, Peter L. Antonelli (ed.), Kluwer Academic Publishers, Dordrecht, 557–966, 2003.

[10] Miron, R. and Tavakol, R., Geometry of space-time and generalized Lagrange spaces, *Publ. Math. Debrecen*, **44**(1–2), 1994, 167–174.

[11] Shen, Z., Differential Geometry of Spray and Finsler Spaces, Kluwer Academic Publishers, Dordrecht, 2001.

[12] Shen, Z., Riemann-Finsler geometry with applications to information geometry, *Chin. Ann. Math.*, **27B**(1), 2006, 73–94.

[13] Shima, H., The Geometry of Hessian Structures, World Scientific Publishing, Hackensack, 2007.

[14] http://en.wikipedia.org/wiki/Normal_distribution