Differential Geometry and its Applications 2 (1992) 197–222
North-Holland

Geometry of statistical manifolds

Mitsunori Noguchi
Nagoya University of Commerce and Business Administration, Sagamine, Aichi, Japan 470–01

Communicated by A. Gray
Received 9 September 1991

Noguchi, M., Geometry of statistical manifolds, Diff. Geom. Appl. 2 (1992) 197–222.

Abstract: A statistical manifold (M, g, ∇) is a Riemannian manifold (M, g) equipped with torsion-free affine connections ∇, ∇^* which are dual with respect to g. A point $p \in M$ is said to be ∇-isotropic if the sectional curvatures have the same value $k(p)$, and (M, g, ∇) is said to be ∇-isotropic when M consists entirely of ∇-isotropic points.

When the difference tensor α of ∇ and the Levi-Civita connection ∇_0 of g is "apolar" with respect to g, Kurose has shown that $\alpha \equiv 0$, and hence $\nabla = \nabla^* = \nabla_0$, provided that $k(p) = k$(constant). His proof relies on the existence of affine immersion which may no longer hold when $k(p)$ is not constant. One objective of this paper is to show that the above Kurose's result still remains valid when (M, g, ∇) is assumed only to be ∇-isotropic. We also discuss the case where (M, g) is complete Riemannian.

Keywords: Statistical manifolds, dual connections, conjugate symmetric statistical manifolds.

MS classification: 53C25, 53A15; 62B10.

1. Introduction

A statistical manifold is, in short, simply a Riemannian manifold (M, g) with one additional structure given by a torsion-free affine connection ∇ and its "dual" connection ∇^*, which is also assumed to be torsion-free; we say ∇ and ∇^* are mutually dual whenever $dg(X, Y) = g(\nabla X, Y) + g(X, \nabla^* Y)$ holds for all vector fields X,Y on M. Thus the geometry of statistical manifold simply reduces to usual Riemannian geometry when ∇ coincides with ∇^*. The notion of dual connection, which is also called conjugate connection in affine geometry, was first introduced into statistics by Amari [1] in his treatment of statistical inference problems and was proven to be quite useful when one deals with certain types of family of probability densities [1]. A statistical manifold is said to be conjugate symmetric [7] whenever the curvature 2-form of ∇ is skew-symmetric with respect to orthonormal frames of (M, g). Even though Lauritzen [7] has shown the existence of a non-conjugate symmetric statistical manifold, the ones which appear often in practice seem mostly to be conjugate symmetric, and from this point of view, geometric characterization of conjugate symmetric statistical manifolds may have some statistical significance. In this paper we deal almost exclusively with

Correspondence to: M. Noguchi, Nagoya University of Commerce and Business Administration, Sagamine, Nisshin-cho, Aichi-gun, Aichi, Japan 470–01. *E-mail*: nogu@jpnnucba

∇-isotropic statistical manifolds, which evidently are included in the set of conjugate symmetric statistical manifolds, and ask under what condition(s) ∇ reduces to the Levi-Civita connection ∇_0 of g.

Meanwhile, Kurose [6] has noticed that there is a close relationship between the geometry of statistical manifold and affine geometry, and has rephrased the work by Calabi [3] and Shima [8] in the language of statistical manifold; this work partly concerns the conditions under which ∇ reduces to the Levi-Civita connection of (M, g).

Section 2 introduces some definitions and notations, and describes the role of difference tensor α, of ∇ and the Levi-Civita connection in our treatment of statistical manifolds. We also discuss what happens when ∇ is required to be left-invariant on some Lie group or required to be almost complex on some Kähler manifold.

In Section 3 we attempt to generalize Kurose's results concerning compact orientable statistical manifolds [6], using Calabi's trick [3] which Shima has mentioned in [8]. We also make use of the decomposition theorem for algebraic curvature tensors, which turns out to be applicable to the curvature tensor of ∇.

Section 4 deals with complete statistical manifolds and derives results similar to the compact case discussed in Section 3. Our proofs in this section rely on a lemma of Cheng-Yau [4].

Section 5 starts out with defining "special" statistical manifold; a statistical manifold is said to be special when $\alpha = \mathrm{Sym}(\psi \otimes g)$ for some 1-form ψ. The above form for α enables us to analyze the behavior of geodesics of ∇, and establishes some conditions under which ∇ becomes geodesically complete, provided (M, g) is complete Riemannian. We also construct some examples of special conjugate symmetric statistical manifolds whose ∇ is geodesically complete. We also remark that the situation above is once again related to Riemannian geometry via conformal Killing fields, as shown in this section.

Finally, I would like to thank my colleague Prof. H. Nagayama for helpful discussions, and my wife Mary for her cooperation during the course of my completing this manuscript.

2 Preliminaries

Throughout this paper we assume that our manifold M is real n-dimensional and connected. Given a C^∞ vector bundle $E \to M$, we define the following objects:

$$\Omega^p(E)(U) = C^\infty E\text{-valued } p\text{-forms over } U,$$

in particular,

$$\Omega^0(E)(U) = C^\infty \text{ sections of } E \text{ over } U,$$

and

$$\Omega^p(E) = \Omega^p(E)(M),$$

$$C^\infty(U) = C^\infty \text{ functions on } U,$$

where U is an open subset of M.

A *connection* ∇ on a C^∞ vector bundle $E \to M$ is a map $\nabla : \Omega^0(E) \to \Omega^1(E)$, satisfying Leibnitz' rule $\nabla(fs) = df \otimes s + f\nabla s$ for all sections $s \in \Omega^0(E)(U)$, $f \in C^\infty(U)$. When $E = TM$ (the tangent bundle of M), $\nabla : \Omega^0(TM) \to \Omega^1(TM)$ is called an *affine connection* on TM. We denote $\Omega^0(TM) = C^\infty$ vector fields on M by $\mathcal{X}M$, and a Riemannian manifold with C^∞ Riemannian metric g by (M, g).

An affine connection ∇^* is called a dual connection of ∇ if

$$dg(X,Y) = g(\nabla X, Y) + g(X, \nabla^* Y)$$

holds for all $X, Y \in \mathcal{X}M$, and as we shall see later, the last identity and ∇ uniquely determines ∇^*, so we speak of *the* dual connection ∇^* of ∇.

A triple (M, g, ∇) is called a *statistical manifold* whenever both ∇ and ∇^* are torsion-free.

Let $e = (e_i)$ be an orthonormal frame on U and $\theta = (\theta^i)$ be the coframe dual to e. In what follows, $\omega_0 = ({\omega_0}^i_j)$ denotes the connection 1-form with respect to e of the Levi-Civita connection ∇_0, and $\omega = (\omega^i_j)$, $\omega^* = ({\omega^*}^i_j)$ denote the connection 1-forms with respect to e of ∇, ∇^*, respectively. We now define the *difference tensor* of ∇ and ∇_0 to be $\alpha = \nabla - \nabla_0 \in \Omega^1(\operatorname{End} TM)$, where $\operatorname{End} TM$ denotes the bundle of endmorphisms of TM.

Here and for the rest of this paper, ^{t} indicates the usual transpose of matrices. By the definition of ∇^*, we have

$$\omega^* + {}^t\omega = 0,$$

and hence

$$\omega^* = -{}^t\omega,$$

which particularly implies that ∇ uniquely determines ∇^*, as mentioned earlier.

We write $\alpha^i{}_j = \alpha(\theta^i \otimes e_j)$ where $\{\theta^i \otimes e_j\}$ is a local frame for $\operatorname{End} TM$ over U, and simply write α for the matrix $(\alpha^i{}_j)$. Substituting

$$\omega = \omega_0 + \alpha$$

into the above expression for ω^*, and using the fact that

$$\omega_0 + {}^t\omega_0 = 0,$$

we obtain

$$\omega^* = \omega_0 - {}^t\alpha. \tag{2.1}$$

Let Θ and Θ^* be the torsion 2-forms of ∇ and ∇^*, respectively, and consider the first structure equations for ∇, ∇^* as follows:

$$\begin{aligned} \Theta &= d\theta + \omega \wedge \theta = d\theta + (\omega_0 + \alpha) \wedge \theta \\ &= (d\theta + \omega_0 \wedge \theta) + \alpha \wedge \theta = \alpha \wedge \theta \end{aligned} \tag{2.2}$$

since $d\theta + \omega_0 \wedge \theta = 0$. Similarly,

$$\begin{aligned}\Theta^* &= d\theta + \omega^* \wedge \theta = d\theta + (\omega_0 - {}^t\alpha) \wedge \theta \\ &= (d\theta + \omega_0 \wedge \theta) - {}^t\alpha \wedge \theta = -{}^t\alpha \wedge \theta.\end{aligned} \tag{2.3}$$

Thus the condition $\Theta = \Theta^* = 0$ is equivalent to requiring that $\alpha \wedge \theta = {}^t\alpha \wedge \theta = 0$. Note that if we write $\alpha^i{}_{jk} = \alpha(\theta^i \otimes e_j)(e_k)$, then $\alpha \wedge \theta = 0$ if and only if $\alpha^i{}_{jk} = \alpha^i{}_{kj}$. These establish the following proposition:

Proposition 2.1. *An affine connection ∇ and its dual connection ∇^* are both torsion-free if and only if $\alpha \wedge \theta = 0$, $\alpha = {}^t\alpha$.*

Proof. Since $\alpha^i{}_{jk} = \alpha^k{}_{ji} = \alpha^k{}_{ij} = \alpha^j{}_{ik}$, $\alpha = {}^t\alpha$. □

Let Ω_0, Ω be the curvature 2-forms of ∇_0, ∇, respectively. Then the second structure equations are as follows:

$$\Omega_0 = d\omega_0 + \omega_0 \wedge \omega_0. \tag{2.4}$$

$$\begin{aligned}\Omega &= d\omega + \omega \wedge \omega = d(\omega_0 + \alpha) + (\omega_0 + \alpha) \wedge (\omega_0 + \alpha) \\ &= d\omega_0 + \omega_0 \wedge \omega_0 + \alpha \wedge \alpha + d\alpha + \omega_0 \wedge \alpha + \alpha \wedge \omega_0 \\ &= \Omega_0 + \alpha \wedge \alpha + d\alpha + \omega_0 \wedge \alpha + \alpha \wedge \omega_0\end{aligned} \tag{2.5}$$

by substituting (2.4).

The Levi-Civita connection ∇_0 induces a *covariant differential*

$$\mathbf{D} : \mathbf{\Omega}^1(\operatorname{End} TM) \to \mathbf{\Omega}^2(\operatorname{End} TM)$$

by $\mathbf{D}\beta = d\beta + \omega_0 \wedge \beta + \beta \wedge \omega_0$ for all $\beta \in \mathbf{\Omega}^1(\operatorname{End} TM)$.

Now we obtain from (2.5) that

$$\Omega = \Omega_0 + \alpha \wedge \alpha + \mathbf{D}\alpha. \tag{2.6}$$

Note that, in terms of $e = (e_1, \ldots, e_n)$, we have a decomposition

$$\Omega = \mathbf{A}(\Omega) + \mathbf{S}(\Omega), \tag{2.7}$$

where $\mathbf{A}(\Omega) = \frac{1}{2}(\Omega - {}^t\Omega)$, $\mathbf{S}(\Omega) = \frac{1}{2}(\Omega + {}^t\Omega)$, and since the decomposition is independent of all choices of an orthonormal frame, it is invariant and also orthogonal. We have $\Omega_0 + {}^t\Omega_0 = 0$ and also

$$\alpha \wedge \alpha + {}^t(\alpha \wedge \alpha) = \alpha \wedge \alpha - {}^t\alpha \wedge {}^t\alpha = \alpha \wedge \alpha - \alpha \wedge \alpha = 0.$$

Furthermore, we compute

$$\begin{aligned}{}^t\mathbf{D}\alpha &= {}^t(d\alpha + \omega_0 \wedge \alpha + \alpha \wedge \omega_0) = d{}^t\alpha - {}^t\alpha \wedge {}^t\omega_0 - {}^t\omega_0 \wedge {}^t\alpha \\ &= d\alpha + \omega_0 \wedge \alpha + \alpha \wedge \omega_0 = \mathbf{D}\alpha.\end{aligned}$$

Combining these facts we deduce

$$\mathbf{A}(\Omega) = \Omega + \alpha \wedge \alpha, \quad \mathbf{S}(\Omega) = \mathbf{D}\alpha. \tag{2.8}$$

A statistical manifold (M, g, ∇) is said to be ∇-*flat* whenever ∇ is a flat affine connection. Let Ω^* be the curvature 2-form of ∇^*. Then, since $\omega^* = -{}^t\omega$, we compute

$$\Omega^* = d\omega^* + \omega^* \wedge \omega^* = -{}^t(d\omega + \omega \wedge \omega),$$

and obtain $\Omega^* = -{}^t\Omega$. Thus, if ∇ is flat, so is ∇^*. Amari [1] has shown that a ∇-flat statistical manifold (M, g, ∇) admits a pair of systems of local coordinates $x = (x^1, \ldots, x^n)$, $y = (y^1, \ldots, y^n)$ on U, which are dual to each other in the following sense: there exists a pair of functions $\varphi, \phi \in C^\infty(U)$, such that $y^i = \partial\varphi/\partial x^i$, $x^i = \partial\phi/\partial y^i$ hold in U. The existence of such local coordinates plays an important role in the theory of statistical inference, as described by Amari [1]. An example of ∇-flat statistical manifold with non-flat Riemannian metric was first constructed on hyperbolic 2-space $\mathbf{H}^2$ by Atkinson, Mitchell [2], Skovgaard [9], and Amari [1] in some detail. Their example can easily be generalized to hyperbolic n-space $\mathbf{H}^n$, as follows:

Example. Let $(x^1, \ldots, x^n)$ be the standard coordinates on $\mathbb{R}^n$ and define

$$\mathbf{H}^n = \{(x^1, \ldots, x^n) \in \mathbb{R}^n : x^n > 0\}.$$

We give $\mathbf{H}^n$ the hyperbolic metric

$$ds^2 = \sum_{i=0}^{n} (dx^i)^2/(x^n)^2,$$

and define an orthonormal coframe on $(\mathbf{H}^n, ds^2)$ by setting

$$\theta^i = dx^i/x^n, \qquad i = 1 \ldots n.$$

The connection 1-form of the Levi-Civita connection ∇_0 is given with respect to $\theta = (\theta^1, \ldots, \theta^n)$ by:

$$\omega = \begin{pmatrix} 0 & \ldots & 0 & -\theta^1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 0 & -\theta^{n-1} \\ \theta^1 & \ldots & \theta^{n-1} & 0 \end{pmatrix},$$

and using the structure equation

$$d\theta = -\omega_0 \wedge \theta,$$

we compute

$$\Omega_0 = d\omega_0 + \omega_0 \wedge \omega_0 = (-\theta^i \wedge \theta^j).$$

Now we define

$$\alpha = \begin{pmatrix} \theta^n & \ldots & 0 & \theta^1 \\ \vdots & \ddots & \ldots & \vdots \\ 0 & \ldots & \theta^n & \theta^{n-1} \\ \theta^1 & \ldots & \theta^{n-1} & 2\theta^n \end{pmatrix}.$$

Then it is an easy exercise to verify that

$$\alpha \wedge \theta = 0, \qquad \alpha = {}^t\alpha,$$
$$\mathbf{A}(\Omega) = \Omega_0 + \alpha \wedge \alpha = 0,$$

and

$$\mathbf{S}(\Omega) = \mathbf{D}\alpha = 0.$$

Hence from (2.7), we conclude $\Omega = 0$ identically on $\mathbf{H}^n$, and the affine connection ∇ determined by $\omega = \omega_0 + \alpha$ turns $(\mathbf{H}^n, ds^2)$ into a desired statistical manifold.

Remark on Kähler manifolds. Let (M, g) be a complex n-dimensional Kähler manifold with its complex structure $\mathbf{J}$ and canonical Kähler connection ∇_0. Suppose that (M, g, ∇) is a statistical manifold. We shall show ∇ coincides with ∇_0, if one imposes on ∇ an additional condition that ∇ is almost complex with respect to $\mathbf{J}$, i.e. $\nabla\mathbf{J} = 0$.

We can find an orthonormal frame $(e_1, \ldots, e_n, \mathbf{J}e_1, \ldots, \mathbf{J}e_n)$ on U, which may be regarded as a unitary frame on U, in the complexified tangent bundle $TM^{\mathbb{C}}$. As is well-known, (see for example [5]), the complex linear extension of $\mathbf{J}$ gives the decomposition $TM^{\mathbb{C}} = T'M + T''M$, where $T'M$, $T''M$ consists of so-called type $(1,0)$ vectors, type $(0,1)$ vectors, respectively. If we define $\epsilon_i = (e_i - \sqrt{-1}\mathbf{J}e_i)/\sqrt{2}$, $\bar{\epsilon}_i = (e_i + \sqrt{-1}\mathbf{J}e_i)/\sqrt{2}$, then ϵ_i, $i = 1, \ldots, n$, span $T'M$ and $\bar{\epsilon}_i, i = 1, \ldots, n$, span $T''M$.

Now, since both ∇_0 and ∇ are almost complex, $\alpha = \omega - \omega_0$ commutes with $\mathbf{J}$ and takes values in the real representation of $gl(n, \mathbb{C}) \subset gl(2n, \mathbb{R})$, where $gl(n, \mathbb{C})$ is the n-dimensional complex general linear algebra and $gl(2n, \mathbb{R})$ the $2n$-dimensional real general linear algebra. Writing $\mathbf{\Omega}^1(\mathbf{g})(U)$ for the set of $\mathbf{g}$-valued 1-forms over U, it follows that α can be represented, with respect to the frame $(\epsilon_1, \ldots, \epsilon_n, \bar{\epsilon}_1, \ldots, \bar{\epsilon}_n)$, as

$$\alpha = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \in \mathbf{\Omega}^1(gl(n, \mathbb{C}))(U) : \; gl(n, \mathbb{C}) \subset gl(2n, \mathbb{R}), \tag{2.9}$$

where $A, B \in \mathbf{\Omega}^1(gl(n, \mathbb{R}))(U)$.

Extending α complex linearly to $TM^{\mathbb{C}}$ and taking $(\epsilon_1, \ldots, \epsilon_n, \bar{\epsilon}_1, \ldots, \bar{\epsilon}_n)$ as a unitary frame, we can represent α as

$$\alpha = \begin{pmatrix} A + \sqrt{-1}B & 0 \\ 0 & A - \sqrt{-1}B \end{pmatrix} \in \mathbf{\Omega}^1(gl(n, \mathbb{C}) \oplus gl(n, \mathbb{C}))(U) : \tag{2.10}$$

$gl(n, \mathbb{C}) \oplus gl(n, \mathbb{C}) \subset gl(2n, \mathbb{C})$, and since $\alpha = {}^t\alpha$ in (2.9) implies $A = {}^tA$, $B = 0$, α in (2.10) will simply be of the form

$$\alpha = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} : A = {}^tA \in \mathbf{\Omega}^1(gl(n, \mathbb{R}))(U). \tag{2.11}$$

Let θ^i, $\theta^{\bar{i}}$ be the forms dual to ϵ_i, $\bar{\epsilon}_i$, respectively. Note that α also must satisfy the condition

$$\alpha \wedge \begin{pmatrix} \theta \\ \bar{\theta} \end{pmatrix} = 0, \tag{2.12}$$

where we write $\theta = (\theta^i)$, $\bar{\theta} = (\theta^{\bar{i}})$.

From (2.11) and (2.12) we obtain the following equation:

$$0 = \alpha \wedge \begin{pmatrix} \theta \\ \bar{\theta} \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \wedge \begin{pmatrix} \theta \\ \bar{\theta} \end{pmatrix} = \begin{pmatrix} A \wedge \theta \\ A \wedge \bar{\theta} \end{pmatrix},$$

or

$$A \wedge \theta = 0, \tag{2.13}$$

$$A \wedge \bar{\theta} = 0. \tag{2.14}$$

Writing $A^i_j = A^i_{jk}\theta^k + A^i_{j\bar{k}}\theta^{\bar{k}}$ and expressing (2.13–14) componentwise, we deduce that A is simultaneously of type $(1,0)$ and of type $(0,1)$, and hence A vanishes identically, from which we conclude that ∇ coincides with ∇_0.

Next we discuss what Lauritzen called *conjugate symmetric* statistical manifolds [7]. A statistical manifold (M, g, ∇) is said to be *conjugate symmetric* whenever the curvature 2-form of ∇ satisfies $\Omega = \mathbf{A}(\Omega)$. From (2.7–8) the last condition is clearly equivalent to $\mathbf{S}(\Omega) = \mathbf{D}\alpha = 0$.

Remark on Lie groups. Let G be a Lie group equipped with bi-invariant metric $\langle \cdot, \cdot \rangle$, and let ∇ be a left invariant affine connection on TG. Suppose $(G, \langle \cdot, \cdot \rangle, \nabla)$ is a conjugate symmetric statistical manifold. We shall now show that, for some Lie groups, ∇ being left invariant is a strong enough condition for concluding that ∇ actually coincides with the Levi-Civita connection ∇_0.

For left invariant vector fields X,Y, and Z, we have $\nabla_{0X}Y = \frac{1}{2}[X,Y]$, and since $\mathbf{D}\alpha = 0$, we obtain

$$\begin{aligned} &[X, \alpha(Y, Z)] - [Y, \alpha(X, Z)] \\ &\qquad - \alpha(Y, [X, Z]) + \alpha(X, [Y, Z]) - 2\alpha([X, Y], Z) = 0, \end{aligned} \tag{2.15}$$

and about the curvature 2-form of ∇, we have

$$\Omega = \Omega_0 + \alpha \wedge \alpha = -\tfrac{1}{4}[[X,Y],Z] + \alpha(X, \alpha(Y,Z)) - \alpha(Y, \alpha(X,Z)), \tag{2.16}$$

following from the fact $\Omega = \mathbf{A}(\Omega)$. Let us choose an orthonormal basis for the lie algebra of left invariant vector fields on G. Let $(C^i{}_{jk})$ denote the structure constants of G, with respect to the basis chosen above. Then we have the following identities for $C^i{}_{jk}$:

$$C^i{}_{jk} = -C^i{}_{kj} = C^j{}_{ki},$$

since $\langle [X,Y], Z\rangle = \langle X, [Y,Z]\rangle$, and

$$C^h{}_{ij}C^l{}_{hk} + C^h{}_{jk}C^l{}_{hi} + C^h{}_{ki}C^l{}_{hj} = 0,$$

which is just the usual Jacobi identity. We next express (2.15) and (2.16) in terms of $(C^i{}_{jk})$ as follows:

$$\begin{aligned} &C^m{}_{il}\alpha^l{}_{jk} - C^m{}_{jl}\alpha^l{}_{ik} - C^l{}_{ik}\alpha^m{}_{jl} + C^l{}_{jk}\alpha^m{}_{il} - 2C^l{}_{ij}\alpha^m{}_{lk} = 0, \\ &\tilde{R}^m{}_{kij} = -\tfrac{1}{4}C^l{}_{ij}C^m{}_{lk} + \alpha^m{}_{il}\alpha^l{}_{jk} - \alpha^m{}_{jl}\alpha^l{}_{ik}, \end{aligned} \tag{2.17}$$

where $\tilde{R}^m{}_{kij}$ are the components of Ω.

As an example we consider $G = SO(3)$ together with the negative of the Killing-Cartan form of it. Then $C^i{}_{jk} = 1$ (-1) whenever (i,j,k) is an even (odd) permutation of $(1,2,3)$, and $C^i{}_{jk} = 0$ otherwise. Thus from (2.17) we conclude that $\alpha^i{}_{jk} = 0$; for example, substituting $i = 1$, $j = 2$, $m = 1$, and $k = 2$ into (2.17), we obtain $\alpha^1{}_{23} = 0$.

Here and for the rest of this paper, we adopt the following notations: $x = (x^1, \ldots, x^n)$ for a system of local coordinates on U, $\partial/\partial x = (\partial/\partial x^1, \ldots, \partial/\partial x^n)$ for a natural frame on U, $\{dx^i \otimes \partial/\partial x^j\}$ for a local frame for $\operatorname{End} TM$ over U; $\bar{\Gamma}^i{}_{jk} \equiv \omega_0(dx^i \otimes \partial/\partial x^j)(\partial/\partial x^k)$, $\Gamma^i{}_{jk} \equiv \omega(dx^i \otimes \partial/\partial x^j)(\partial/\partial x^k)$, $\alpha^i{}_{jk} \equiv \alpha(dx^i \otimes \partial/\partial x^j)(\partial/\partial x^k)$, $\bar{\Gamma}_{ijk} \equiv g_{il}\bar{\Gamma}^l{}_{jk}$, $\Gamma_{ijk} \equiv g_{il}\Gamma^l{}_{jk}$, $\alpha_{ijk} \equiv g_{il}\alpha^l{}_{jk}$, where (g_{ij}) are the components of g with respect to $\partial/\partial x$. For $\beta \in \mathbf{\Omega}^1(\operatorname{End} TM)$, the components of $\mathbf{D}\beta$ take the form $(\mathbf{D}\beta)^i{}_{jkl} = \beta^i{}_{jk,l} - \beta^i{}_{jl,k}$, where the index of the Levi-Civita covariant differentiation is written to the right, following the comma.

Remark. From Proposition 2.1 it follows that given an affine connection ∇ on (M, g), (M, g, ∇) becomes a statistical manifold if and only if the components of the difference tensor α_{ijk} are symmetric in all pairs of indices.

We now define $\operatorname{Tr}\nabla \equiv \operatorname{Tr}\alpha$, where $\operatorname{Tr}\alpha$ is taken in the endomorphism part, and remark that this $\operatorname{Tr}\alpha$ indeed is a well-defined 1-form on M, whose components in terms of local coordinates are given as $(\operatorname{Tr}\alpha)_k = \alpha^j{}_{jk} = g^{ij}\alpha_{ijk}$, where (g^{ij}) is the inverse of the matrix (g_{ij}) of the metric g.

It is interesting to note that for certain classes of statistical manifolds, a condition $\operatorname{Tr}\nabla \equiv 0$ is sufficient for ∇ to be reduced to the Levi-Civita connection ∇_0. An example of such classes of statistical manifolds is provided by a theorem of Calabi [3], which is restated in the paper by Shima [8] in conjunction with locally Hessian structure. The theorem mentioned above can be restated in the language of statistical manifold as follows:

Theorem 2.2 (Calabi [3]). *Let (M, g, ∇) be a statistical manifold such that M is an open domain in the n-dimensional real affine space and ∇ the natural flat affine connection. If (M, g) is complete Riemannian and if* $\operatorname{Tr}\nabla \equiv 0$ *on M, then ∇ and ∇_0 coincide (and of course g is flat).*

3. Compact conjugate symmetric statistical manifolds

For the rest of this paper we adopt the following notations: $\beta \equiv \nabla \operatorname{Tr}\nabla$, $\operatorname{Tr}\beta \equiv g^{ij}\beta_{ij}$, $\|\cdot\|$ for the Riemannian length of what is inside, $R^i{}_{jkl}$ for the components of Ω_0, $\tilde{R}^i{}_{jkl}$ for the components of Ω. Moreover, we shall follow the usual convention of raising or lowering indices of tensors and also the usual convention of contracting tensors. We shall write, as for the Riemannian curvature tensor, $\tilde{R}_{ijkl} = g_{im}\tilde{R}^m{}_{jkl}$, $\tilde{R}_{jl} = g^{ik}\tilde{R}_{ijkl}$, and $\tilde{R} = g^{ij}\tilde{R}_{ij}$.

We now attempt to find some conditions on a compact conjugate symmetric statistical manifold (M, g, ∇) under which the affine connection ∇ reduces to the Levi-Civita connection ∇_0. Since a ∇ -flat statistical manifold is clearly conjugate symmetric, the following theorem of Shima is interesting from the above point of view and can be rephrased to read as:

Theorem 3.1 (Shima [8]). *Let (M, g, ∇) be a compact orientable ∇-flat statistical manifold. Then:*

(1) $$\int_M \operatorname{Tr}\beta = \int_M \|\operatorname{Tr}\nabla\|^2 \geqslant 0;$$

(2) *If $\int_M \operatorname{Tr}\beta = 0$, then ∇ coincides with ∇_0.*

Remark. By definition $\operatorname{Tr}\nabla = \operatorname{Tr}\alpha \in \Omega^1(M)$. Shima has noticed that $\nabla \operatorname{Tr}\alpha = \nabla_0 \operatorname{Tr}\alpha - \alpha \operatorname{Tr}\alpha$, from which $\operatorname{Tr}(\nabla \operatorname{Tr}\alpha) = \operatorname{Tr}(\nabla_0 \operatorname{Tr}\alpha) - \|\operatorname{Tr}\alpha\|^2$ follows. We now integrate the last expression over M to establish (1). He has also noticed that a trick devised by Calabi [3] works for proving (2).

∇ is said to have *constant curvature* k [6] whenever the curvature tensor satisfies

$$\tilde{R}_{ijkl} = \frac{k}{n(n-1)}(g_{ik}g_{jl} - g_{il}g_{jk}),$$

or equivalently

$$\tilde{\Omega}^i_j = k\, \theta^i \wedge \theta^j,$$

where $\theta = (\theta^i)$ is an orthonormal coframe. It is clear that a statistical manifold with ∇ having constant curvature is conjugate symmetric.

Remark. Schur's theorem ($n \geqslant 3$) does not hold in general, since the covariant derivative by ∇ of the metric tensor may not vanish.

We now state the following theorem due to Kurose, which slightly generalizes Theorem 3.1.

Theorem 3.2 (Kurose [6]). *Let (M, g, ∇) be a compact orientable statistical manifold with ∇ having constant curvature k. Then:*

(1) $$\int_M \operatorname{Tr}\beta = \int_M \|\operatorname{Tr}\nabla\|^2 \geqslant 0;$$

(2) *If $\int_M \operatorname{Tr}\beta = 0$ and if $k \geqslant 0$, then ∇ coincides with ∇_0.*

Remark. $\beta = \nabla \operatorname{Tr}\nabla$ has the local expression $\beta_{ij} = \alpha_{i,j} - \alpha^k{}_{ij}\alpha^l{}_{lk}$ and since (M, g, ∇) is conjugate symmetric, we have by definition $\mathbf{D}\alpha = \alpha_{ijk,l} - \alpha_{ijl,k} = 0$ and hence have $\alpha_{i,j} = \alpha_{j,i}$, which indicates β is symmetric. We can now rephrase the conclusions of Theorem 3.2 to read:

(1) β cannot be negative definite.
(2) if β is negative semi-definite and if $k \geqslant 0$, then ∇ coincides with ∇_0.

The last two statements are identical to the conclusions of Corollary 2 in [6].

The following theorem shows that on 2-sphere $\mathbf{S}^2$ with its standard metric g_0, Theorem 3.2 has a further generalization.

Theorem 3.3. *Let $(\mathbf{S}^2, g_0, \nabla)$ be a conjugate symmetric statistical manifold. Then:*

(1) $$\int_{\mathbf{S}^2} \operatorname{Tr}\beta = \int_{\mathbf{S}^2} \|\operatorname{Tr}\nabla\|^2 \geqslant 0$$

(2) *If $\int_{\mathbf{S}^2} \operatorname{Tr}\beta = 0$, then ∇ coincides with ∇_0.*

Proof. We only need to prove (2). Since $(\mathbf{S}^2, g_0, \nabla)$ is conjugate symmetric we have

$$\mathbf{D}\alpha = 0, \tag{3.1}$$

or

$$\alpha^i{}_{jk,l} = \alpha^i{}_{jl,k}.$$

We may choose a system of local coordinates (x, y) on $U \subset \mathbf{S}^2$, identifying U with $\mathbb{R}^2$ and expressing the metric $ds^2 = g_0$ as

$$ds^2 = \frac{4(dx^2 + dy^2)}{(1 + x^2 + y^2)^2}.$$

Since α must satisfy

$$\alpha \wedge \begin{pmatrix} dx \\ dy \end{pmatrix} = 0, \quad \alpha = {}^t\alpha, \quad \text{and} \quad \operatorname{Tr}\nabla = \operatorname{Tr}\alpha = 0,$$

we have

$$\alpha = \begin{pmatrix} Bdx + Ady & Adx - Bdy \\ Adx - Bdy & -Bdx - Ady \end{pmatrix}$$

for some $A,\ B \in C^\infty(\mathbb{R}^2)$. Then (3.1) is equivalent to

$$\begin{aligned} \frac{\partial A}{\partial x} - \frac{\partial B}{\partial y} &= \frac{4}{1 + x^2 + y^2}(Ax - By), \\ \frac{\partial B}{\partial x} + \frac{\partial A}{\partial y} &= \frac{4}{1 + x^2 + y^2}(Bx + Ay). \end{aligned} \tag{3.2}$$

Introducing a complex coordinate $z = x + \sqrt{-1}y$ on $\mathbb{R}^2$, we define

$$\frac{\partial}{\partial z} = \frac{1}{2}\Big(\frac{\partial}{\partial x} - \sqrt{-1}\frac{\partial}{\partial y}\Big), \quad \frac{\partial}{\partial \bar z} = \frac{1}{2}\Big(\frac{\partial}{\partial x} + \sqrt{-1}\frac{\partial}{\partial y}\Big),$$

where "$^-$" indicates the usual complex conjugation. Setting $f(z, \bar z) = A + \sqrt{-1}B$, we can now express (3.2) as:

$$\frac{\partial f}{\partial \bar z} - \frac{2z}{1 + |z|^2} f = 0, \tag{3.3}$$

where $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^2$. Introducing another C^∞ complex valued function $g(z,\bar{z})$ by requiring $f = (1+|z|^2)g$, and substituting the last expression into (3.3), we obtain

$$\frac{\partial f}{\partial \bar{z}} - \frac{2z}{1+|z|^2} f = (1+|z|^2)^2 \frac{\partial g}{\partial \bar{z}}.$$

Thus f satisfies (3.3) if and only if g is holomorphic. Moreover, since $|\alpha|^2_{g_0} = (1+|z|^2)^2(A^2+B^2) = (1+|z|^2)^2|f|^2 = (1+|z|^2)^6|g|^2$, regularity at $z=\infty$ and Liouville's theorem imply $g \equiv 0$, and hence $\alpha \equiv 0$. □

For a general compact manifold other than $\mathbf{S}^2$, we still can slightly generalize Theorem 3.2 using Calabi's trick, which was also used by Shima in proving Theorem 3.1.

Theorem 3.4. *Let (M,g,∇) be a compact orientable statistical manifold with affine connection ∇ whose curvature tensor satisfies*

$$\tilde{R}_{ijkl} = \frac{\tilde{R}}{n(n-1)}(g_{ik}g_{jl} - g_{il}g_{jk}).$$

Then:

(1) $$\int_M \operatorname{Tr}\beta = \int_M \|\operatorname{Tr}\nabla\|^2 \geqslant 0.$$

(2) *If $\int_M \operatorname{Tr}\beta = 0$, and if $\int_M \tilde{R}(R-\tilde{R}) \geqslant 0$, then ∇ coincides with ∇_0. Consequently, if $n \geqslant 3$, Shur's theorem implies (M,g) is a space of constant curvature $\tilde{R}/(n(n-1))$ with constant $\tilde{R} \geqslant 0$.*

Remark. Here $\tilde{R}$ is not assumed to be constant. Moreover, as we shall see in the following proof, $\int_M \operatorname{Tr}\beta = 0$ implies $R - \tilde{R} = \|\alpha\|^2 \geqslant 0$, and hence the condition $\int_M \tilde{R}(R-\tilde{R}) \geqslant 0$ is weaker than $\tilde{R} \geqslant 0$.

Proof. We only need to prove (2). Since (M,g,∇) is conjugate symmetric, we have $\mathbf{D}\alpha = 0$ or $\alpha_{ijk,l} = \alpha_{ijl,k}$; observe that the last condition is equivalent to saying $\alpha_{ijk,l}$ are symmetric in all pairs of indices. We also have $\Omega = \Omega_0 + \alpha\wedge\alpha$, which can be written in terms of components as

$$R_{ijkl} - \tilde{R}_{ijkl} = \alpha_{il}{}^m\alpha_{jkm} - \alpha_{ik}{}^m\alpha_{jlm}. \tag{3.4}$$

From (3.4) we obtain

$$R_{ij} - \tilde{R}_{ij} = \alpha_i{}^{kl}\alpha_{jkl} - \alpha^k\alpha_{ijk}, \tag{3.5}$$

where we simply write $\alpha_k = \alpha^i{}_{ik}$, and contracting (3.5) we get

$$\begin{aligned} R - \tilde{R} &= \alpha^{ijk}\alpha_{ijk} - \alpha^i\alpha_i \\ &= \|\alpha_{ijk}\|^2 - \|\alpha_i\|^2. \end{aligned} \tag{3.6}$$

Let Δ be the Laplace operator of g. We then have from (3.6):

$$\tfrac{1}{2}\Delta(R-\tilde{R}) = \tfrac{1}{2}(\alpha^{ijk}\alpha_{ijk})_{,l}{}^{l} - \tfrac{1}{2}\Delta\|\alpha_i\|^2, \tag{3.7}$$

and computing the first term on the right, we obtain

$$\begin{aligned}\tfrac{1}{2}(\alpha^{ijk}\alpha_{ijk})_{,l}{}^{l} &= (\alpha^{ijk}\alpha_{ijk,l})_{,}{}^{l}\\ &= \|\alpha_{ijk,l}\|^2 + \alpha^{ijk}\alpha_{ijk,l}{}^{l}.\end{aligned} \tag{3.8}$$

We now compute

$$\begin{aligned}\alpha^{ijk}\alpha_{ijk,l}{}^{l} &= \alpha^{ijk}g^{lm}\alpha_{ijk,lm}\\ &= \alpha^{ijk}g^{lm}\alpha_{ljk,im} \quad (\text{Note that } \alpha_{ijk,l} = \alpha_{ijl,k})\\ &= \alpha^{ijk}g^{lm}(\alpha_{ljk,mi} + \alpha_{sjk}R^{s}{}_{lim} + \alpha_{lsk}R^{s}{}_{jim} + \alpha_{ljs}R^{s}{}_{kim})\\ &= \alpha^{ijk}\alpha_{j,ki} + R_{ij}\alpha^{ikl}\alpha^{j}{}_{kl} + 2R_{ijkl}\alpha^{ilm}\alpha^{jk}{}_{m}.\end{aligned} \tag{3.9}$$

Remark. The trick we have used above to get from the second step to the third step and then from the third step to the fourth step is due to Calabi [3] and is also used by Shima [8].

Substituting (3.9) into (3.8) we get

$$\tfrac{1}{2}(\alpha^{ijk}\alpha_{ijk})_{,l}{}^{l} = \|\alpha_{ijk,l}\|^2 + \alpha^{ijk}\alpha_{j,ki} + R_{ij}\alpha^{ikl}\alpha^{j}{}_{kl} + 2R_{ijkl}\alpha^{ilm}\alpha^{jk}{}_{m}. \tag{3.10}$$

We then substitute (3.10) into (3.7) to establish

$$\begin{aligned}\tfrac{1}{2}\Delta(R-\tilde{R}) &= R_{ij}\alpha^{ikl}\alpha^{j}{}_{kl} + 2R_{ijkl}\alpha^{ilm}\alpha^{jk}{}_{m}\\ &\quad + \|\alpha_{ijk,l}\|^2 + \alpha^{ijk}\alpha_{j,ki} - \tfrac{1}{2}\Delta\|\alpha_i\|^2\\ &= (R_{ij}-\tilde{R}_{ij})\alpha^{ikl}\alpha^{j}{}_{kl} + 2(R_{ijkl}-\tilde{R}_{ijkl})\alpha^{ilm}\alpha^{jk}{}_{m}\\ &\quad + \tilde{R}_{ij}\alpha^{ikl}\alpha^{j}{}_{kl} + 2\tilde{R}_{ijkl}\alpha^{ilm}\alpha^{jk}{}_{m}\\ &\quad + \|\alpha_{ijk,l}\|^2 + \alpha^{ijk}\alpha_{j,ki} - \tfrac{1}{2}\Delta\|\alpha_i\|^2.\end{aligned} \tag{3.11}$$

Observe that from (3.4) we have

$$\begin{aligned}\|R_{ijkl}-\tilde{R}_{ijkl}\|^2 &= (R_{ijkl}-\tilde{R}_{ijkl})\alpha^{ilm}\alpha^{jk}{}_{m} - (R_{ijkl}-\tilde{R}_{ijkl})\alpha^{ikm}\alpha^{jl}{}_{m}\\ &= 2(R_{ijkl}-\tilde{R}_{ijkl})\alpha^{ilm}\alpha^{jk}{}_{m},\end{aligned} \tag{3.12}$$

and from (3.5) we have

$$\|R_{ij}-\tilde{R}_{ij}\|^2 = (R_{ij}-\tilde{R}_{ij})\alpha^{ikl}\alpha^{j}{}_{kl} - (R_{ij}-\tilde{R}_{ij})\alpha^{k}\alpha_{k}{}^{ij}.$$

Combining (3.11) through (3.13) we obtain

$$\begin{aligned}\tfrac{1}{2}\Delta(R-\tilde{R}) &= \|R_{ij}-\tilde{R}_{ij}\|^2 + \|R_{ijkl}-\tilde{R}_{ijkl}\|^2 + \|\alpha_{ijk,l}\|^2\\ &\quad + \tilde{R}_{ij}\alpha^{ikl}\alpha^{j}{}_{kl} + 2\tilde{R}_{ijkl}\alpha^{ilm}\alpha^{jk}{}_{m}\\ &\quad + (R_{ij}-\tilde{R}_{ij})\alpha^{k}\alpha_{k}{}^{ij} + \alpha^{ijk}\alpha_{j,ki} - \tfrac{1}{2}\Delta\|\alpha_i\|^2.\end{aligned} \tag{3.14}$$

We now observe that if $\int_M \operatorname{Tr}\beta = 0$, then $\operatorname{Tr}\nabla \equiv 0$ or $\alpha_i \equiv 0$, and substituting the last expression into (3.14) we establish

$$\begin{aligned}\tfrac{1}{2}\Delta(R-\tilde{R}) &= \|R_{ij}-\tilde{R}_{ij}\|^2 + \|R_{ijkl}-\tilde{R}_{ijkl}\|^2 + \|\alpha_{ijk,l}\|^2 \\ &\quad + \tilde{R}_{ij}\alpha^{ikl}\alpha^{j}{}_{kl} + 2\tilde{R}_{ijkl}\alpha^{ilm}\alpha^{jk}{}_{m}.\end{aligned} \tag{3.15}$$

If

$$\tilde{R}_{ijkl} = \frac{\tilde{R}}{n(n-1)}(g_{ik}g_{jl} - g_{il}g_{jk}),$$

then (3.15) becomes

$$\begin{aligned}\tfrac{1}{2}\Delta(R-\tilde{R}) &= \|R_{ij}-\tilde{R}_{ij}\|^2 + \|R_{ijkl}-\tilde{R}_{ijkl}\|^2 + \|\alpha_{ijk,l}\|^2 \\ &\quad + \frac{n+1}{n(n-1)}\tilde{R}\|\alpha_{ijk}\|^2.\end{aligned} \tag{3.16}$$

Since $\alpha_i \equiv 0$, (3.6) reduces to $R - \tilde{R} = \|\alpha_{ijk}\|^2$, and substituting the last expression into (3.16), we obtain

$$\begin{aligned}\tfrac{1}{2}\Delta(R-\tilde{R}) &= \|R_{ij}-\tilde{R}_{ij}\|^2 + \|R_{ijkl}-\tilde{R}_{ijkl}\|^2 + \|\alpha_{ijk,l}\|^2 \\ &\quad + \frac{n+1}{n(n-1)}\tilde{R}(R-\tilde{R}).\end{aligned} \tag{3.17}$$

Integrating both sides of (3.17) over M, and remembering $\int_M \tilde{R}(R-\tilde{R}) \geqslant 0$, we deduce that $R_{ijkl} = \tilde{R}_{ijkl}$, from which we obtain $R = \tilde{R}$ and conclude that α vanishes identically on M. □

It is well-known that a Riemannian curvature tensor R_{ijkl}, as an "algebraic curvature tensor," decomposes orthogonally into the following three parts: Weyl part, trace-free Ricci part, and scalar curvature part. Observe that the only Riemannian properties needed for the decomposition are: (1) the endomorphism part of curvature 2-forms is skew-symmetric with respect to orthonormal frames; (2) connections are torsion-free; (3) Bianchi's 1st identity holds. For a conjugate symmetric statistical manifold we note that ∇ and its curvature 2-form Ω have all the properties listed above, and hence the decomposition can be applied to $\tilde{R}_{ijkl}$ and also to the difference $(R_{ijkl} - \tilde{R}_{ijkl})$. Consequently, we have

$$\|R_{ijkl}-\tilde{R}_{ijkl}\|^2 = \begin{cases} \|W_{ijkl}-\tilde{W}_{ijkl}\|^2 + \dfrac{4}{n-2}\|G_{ij}-\tilde{G}_{ij}\|^2 \\ \qquad + \dfrac{2}{n(n-1)}(R-\tilde{R})^2 \quad \text{for } n \geqslant 4, \\ 4\|G_{ij}-\tilde{G}_{ij}\|^2 + \tfrac{2}{3}(R-\tilde{R})^2 \quad \text{for } n = 3, \\ (R-\tilde{R})^2 \quad \text{for } n = 2, \end{cases} \tag{3.18}$$

where W_{ijkl}, $\tilde{W}_{ijkl}$ correspond to the Weyl part, and G_{ij}, $\tilde{G}_{ij}$ correspond to the trace-free Ricci part in the above decomposition.

We now suppose that $\operatorname{Tr}\nabla \equiv 0$ (so that $\alpha_i \equiv 0$) and

$$\tilde{R}_{ijkl} = \frac{\tilde{R}}{n(n-1)}(g_{ik}g_{jl} - g_{il}g_{jk});$$

the last expression clearly implies $\tilde{W}_{ijkl} = \tilde{G}_{ij} = 0$ at all points in M.

Combining (3.17) and (3.18), we obtain

$$\tfrac{1}{2}\Delta(R-\tilde{R}) = \begin{cases} \|R_{ij} - \tilde{R}_{ij}\|^2 + \|W_{ijkl}\|^2 + \dfrac{4}{n-2}\|G_{ij}\|^2 \\ \quad + \dfrac{2}{n(n-1)}(R-\tilde{R})^2 + \|\alpha_{ijk,l}\|^2 \\ \quad + \dfrac{n+1}{n(n-1)}\tilde{R}(R-\tilde{R}) \quad \text{for } n \geqslant 4, \\ \|R_{ij} - \tilde{R}_{ij}\|^2 + 4\|G_{ij}\|^2 + \frac{2}{3}(R-\tilde{R})^2 + \|\alpha_{ijk,l}\|^2 \\ \quad + \frac{2}{3}\tilde{R}(R-\tilde{R}) \quad \text{for } n = 3, \\ \|R_{ij} - \tilde{R}_{ij}\|^2 + (R-\tilde{R})^2 + \|\alpha_{ijk,l}\|^2 \\ \quad + \frac{3}{2}\tilde{R}(R-\tilde{R}) \quad \text{for } n = 2. \end{cases} \tag{3.19}$$

We observe that (3.19), together with obvious inequality $\|R_{ij} - \tilde{R}_{ij}\|^2 \geqslant (1/n)(R-\tilde{R})^2$, (where equality holds when $n = 2$), establishes

$$\Delta(R-\tilde{R}) \geqslant \begin{cases} 2\|W_{ijkl}\|^2 + \dfrac{8}{n-2}\|G_{ij}\|^2 \\ \quad + \dfrac{2(n+1)}{n(n-1)}(R-\tilde{R})R \quad \text{for } n \geqslant 4, \\ 8\|G_{ij}\|^2 + \frac{4}{3}(R-\tilde{R})R \quad \text{for } n = 3, \\ 3(R-\tilde{R})R \quad \text{for } n = 2. \end{cases} \tag{3.20}$$

As an application of what we have just derived we prove the following theorem which can be thought of as generalization of Theorem 3.3:

Theorem 3.5. *Let (M, g, ∇) be a compact orientable statistical manifold with non-negative scalar curvature R and affine connection ∇ whose curvature tensor satisfies*

$$\tilde{R}_{ijkl} = \frac{\tilde{R}}{n(n-1)}(g_{ik}g_{jl} - g_{il}g_{jk}).$$

Then:

(1) $$\int_M \operatorname{Tr}\beta = \int_M \|\operatorname{Tr}\nabla\|^2 \geqslant 0.$$

(2) *If* $\int_M \operatorname{Tr}\beta = 0$, *then* $R_{ijkl} \equiv 0$ *and* $\tilde{R} = \text{const} < 0$, *unless* $\nabla = \nabla_0$ (*and* $R = \tilde{R} = \text{const}$ *for* $n \geqslant 3$).

Proof. We shall prove (2) only for the case $n \geqslant 4$, since the case $n = 2, 3$ can be treated similarly. We have from (3.20) the inequality $\Delta(R - \tilde{R}) \geqslant 0$ due to the fact that $R \geqslant 0$. Observing that $R - \tilde{R} = \|\alpha_{ijk}\|^2 \geqslant 0$, we obtain by Bochner's lemma that $R - \tilde{R} = \text{const} \geqslant 0$. If $R - \tilde{R} = \|\alpha_{ijk}\|^2 = 0$, then $\nabla = \nabla_0$, and by Schur's theorem it follows that $R = \tilde{R} = \text{const} \geqslant 0$. Observe that (3.20) becomes

$$0 \geqslant 2\|W_{ijkl}\|^2 + \frac{8}{n-2}\|G_{ij}\|^2 \frac{2(n+1)}{n(n-1)}(R-\tilde{R})R.$$

Thus if $R - \tilde{R} = \|\alpha_{ijk}\|^2 > 0$, then we must have $W_{ijkl} = G_{ij} = R = 0$ at all points in M. But this implies $R_{ijkl} \equiv 0$, hence $R \equiv 0$, and therefore $\tilde{R} = \text{const} < 0$. □

4. Complete Riemannian conjugate symmetric statistical manifolds

Let (M, g, ∇) be a conjugate symmetric statistical manifold such that (M, g) is complete Riemannian. Here we shall prove the following theorem which is analogous to Theorem 3.5 for compact case:

Theorem 4.1. *Let (M, g, ∇) be a statistical manifold such that (M, g) is complete Riemannian. Suppose* $\operatorname{Tr}\nabla \equiv 0$, $\tilde{R} \geqslant 0$, *and*

$$\tilde{R}_{ijkl} = \frac{\tilde{R}}{n(n-1)}(g_{ik}g_{jl} - g_{il}g_{jk}).$$

Then ∇ coincides with ∇_0.

Proof. We have from (3.20) the following inequality which holds for all dimensions greater than one:

$$\Delta(R-\tilde{R}) \geqslant \frac{2(n-1)}{n(n-1)}(R-\tilde{R})R. \tag{4.1}$$

Remembering $\tilde{R} \geqslant 0$ and $R - \tilde{R} = \|\alpha_{ijk}\|^2 \geqslant 0$, we compute the term on the right of (4.1) as

$$\begin{aligned}\frac{2(n+1)}{n(n-1)}(R-\tilde{R})R &= \frac{2(n+1)}{n(n-1)}(R-\tilde{R})^2 + \frac{2(n+1)}{n(n-1)}\tilde{R}(R-\tilde{R})\\ &\geqslant \frac{2(n+1)}{n(n-1)}(R-\tilde{R})^2,\end{aligned}$$

and combining this with (4.1), we establish

$$\Delta(R-\tilde{R}) \geqslant \frac{2(n+1)}{n(n-1)}(R-\tilde{R})^2. \tag{4.2}$$

We now observe that from (3.5) the Ricci tensor R_{ij} satisfies $R_{ij} - \tilde{R}_{ij} = \alpha_i^{kl}\alpha_{jkl}$ since $\alpha_i \equiv 0$, and this implies that $(R_{ij} - \tilde{R}_{ij})$ is positive semi-definite, and hence R_{ij} is bounded from below since $\tilde{R}_{ij} = \tilde{R}g_{ij}/n$, where $\tilde{R} \geqslant 0$ by assumption. Now we can apply [4, Corollary of Theorem 8], derived by Cheng and Yau, to the inequality (4.2) to conclude that $R - \tilde{R} = \|\alpha_{ijk}\|^2 = 0$ at all points in M and hence that $\nabla = \nabla_0$. □

Remark. Theorem 4.1 may be thought of as a generalization of Theorem 2.2 in the sense that M is no longer required to be an open domain in $\mathbb{R}^n$ and ∇ is allowed to have non-zero curvature.

We also remark that if $\tilde{R}$ is bounded away from zero, then R_{ij} will be strictly positive, and Bonnet's theorem implies that M is in fact compact, since (M, g) is assumed to be complete Riemannian.

We again make use of Corollary of Theorem 8 of Cheng and Yau [4] to establish the following theorem:

Theorem 4.2. *Let (M, g, ∇) be a statistical manifold such that (M, g) is complete Riemannian. Suppose* $\mathrm{Tr}\,\nabla \equiv 0$, $R \geqslant \epsilon$ *for some constant* $\epsilon > 0$, $\inf_M \tilde{R}$ *exists, and*

$$\tilde{R}_{ijkl} = \frac{\tilde{R}}{n(n-1)}(g_{ik}g_{jl} - g_{il}g_{jk}).$$

Then ∇ coincides with ∇_0.

Proof. Choose a constant α such that $0 < \alpha < 1$. Since $R \geqslant \epsilon > 0$, we have

$$R - (\tfrac{1}{2}\epsilon)^{1-\frac{1}{\alpha}} R^{\frac{1}{\alpha}} \leqslant \epsilon - (\tfrac{1}{2}\epsilon)^{1-\frac{1}{\alpha}} \epsilon^{\frac{1}{\alpha}} < 0$$

at all points in M. Now, since $\inf_M \tilde{R}$ is assumed to exist, we can choose a constant $c > 0$ large enough to ensure

$$\tilde{R} \geqslant c\{\epsilon - (\tfrac{1}{2}\epsilon)^{1-\frac{1}{\alpha}} \epsilon^{\frac{1}{\alpha}}\} \geqslant c\{R - (\tfrac{1}{2}\epsilon)^{1-\frac{1}{\alpha}} R^{\frac{1}{\alpha}}\}$$

at all points in M. We then have

$$(\tfrac{1}{2}\epsilon)^{1-\frac{1}{\alpha}} R^{\frac{1}{\alpha}} \geqslant R - \frac{\tilde{R}}{c} \geqslant \frac{1}{c}(R - \tilde{R}), \tag{4.3}$$

taking $c \geqslant 1$ if necessary.

From (4.3) we obtain

$$R \geqslant \frac{1}{c^{\alpha}}\left(\frac{2}{\epsilon}\right)^{\alpha-1}(R - \tilde{R})^{\alpha},$$

and combining this with (4.1), we establish

$$\begin{aligned}\Delta(R - \tilde{R}) &\geqslant \frac{2(n+1)}{n(n-1)}(R - \tilde{R})R \\ &\geqslant \frac{2(n+1)}{n(n-1)}\frac{1}{c^{\alpha}}\left(\frac{2}{\epsilon}\right)^{\alpha-1}(R - \tilde{R})^{1+\alpha}.\end{aligned}$$

As in the proof of Theorem 4.1, $(R_{ij} - \tilde{R}_{ij})$ is positive semi-definite, and since $\inf_M \tilde{R}$ exists, it follows that R_{ij} is bounded from below. Now [4, Corollary of Theorem 8] is applicable, and we deduce that $\|\alpha_{ijk}\|^2 = R - \tilde{R} = 0$ at all points in M, and hence $\nabla = \nabla_0$. □

Example. We consider $(\mathbb{R}^2, ds^2, \nabla)$, where $ds^2 = dx^2 + dy^2$ is the standard metric on $\mathbb{R}^2$, in the natural coordinates $(x^1, x^2) = (x, y)$. Note that $\tilde{\Omega}^1_2 = \frac{1}{2}\tilde{R}\, dx \wedge dy$. If we assume $\operatorname{Tr} \nabla \equiv 0$, then α can be expressed as

$$\alpha = \begin{pmatrix} dB & dA \\ dA & -dB \end{pmatrix},$$

where A, $B \in C^\infty(\mathbb{R}^2)$. Note that the condition

$$\alpha \wedge \begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} dB & dA \\ dA & -dB \end{pmatrix} \wedge \begin{pmatrix} dx \\ dy \end{pmatrix} = 0$$

is equivalent to requiring that the function $f = A + \sqrt{-1}B$ is holomorphic on $\mathbb{R}^2$. Moreover, since $\Omega_0 \equiv 0$, we have

$$\tilde{\Omega} = \alpha \wedge \alpha = \begin{pmatrix} 0 & -2dA \wedge dB \\ 2dA \wedge dB & 0 \end{pmatrix},$$

from which we obtain

$$\tilde{R} = -4\left\{\left(\frac{\partial A}{\partial x}\right)^2 + \left(\frac{\partial A}{\partial y}\right)^2\right\} = -4\left\{\left(\frac{\partial B}{\partial x}\right)^2 + \left(\frac{\partial B}{\partial y}\right)^2\right\}.$$

The above example shows that the conclusion of Theorem 4.1 no longer holds if we drop the condition $\tilde{R} \geqslant 0$. It also shows that in Theorem 4.2 the condition $R \geqslant \epsilon > 0$ is essential and the condition that $\inf_M \tilde{R}$ exists is also essential; to see this take, for example, $f = (x + \sqrt{-1}y)^2$.

If we set $f = x + \sqrt{-1}y$, then

$$\alpha = \begin{pmatrix} dy & dx \\ dx & -dy \end{pmatrix},$$

and we compute $\tilde{R} = -4$. Observe now that the whole situation descends to 2-torus $\mathbf{T}^2$ via a covering projection $\mathbb{R}^2 \to \mathbf{T}^2$. This shows that the condition $\int_M \tilde{R}(R - \tilde{R}) \geqslant 0$ in Theorem 3.4 is essential. However, we do not know whether the condition $R \geqslant 0$ in Theorem 3.5 is essential or not.

5. Special conjugate symmetric statistical manifolds

Let (M, g, ∇) be a conjugate symmetric statistical manifold so that $\alpha_{ijk,l}$ are symmetric in all pairs of indices (since $\mathbf{D}\alpha = 0$ identically on M). A statistical manifold (M, g, ∇) is said to be *special* whenever we can write

$$\alpha = \operatorname{Sym}(\psi \otimes g)$$

for some $\psi \in \mathbf{\Omega}^1(M)$, where $\operatorname{Sym}(\psi \otimes g)$ means the usual symmetrization of $\psi \otimes g$.

We now compute the following quantities:

$$\alpha_{ijk} = \tfrac{1}{3}(\psi_i g_{jk} + \psi_j g_{ki} + \psi_k g_{ij}), \tag{5.1}$$

$$\alpha_k = g^{ij}\alpha_{ijk} = \tfrac{1}{3}g^{ij}(\psi_i g_{jk} + \psi_j g_{ki} + \psi_k g_{ij}) = \tfrac{1}{3}(n+2)\psi_k, \tag{5.2}$$

$$\|\alpha_{ijk}\|^2 = \alpha^{ijk}\alpha_{ijk} = \tfrac{1}{3}(n+2)\|\psi_i\|^2, \tag{5.3}$$

$$\|\alpha_i\|^2 = \tfrac{1}{9}(n+2)^2\|\psi_i\|^2, \tag{5.4}$$

$$\begin{aligned} R_{ijkl} - \tilde{R}_{ijkl} = -\tfrac{1}{9}\{&\|\psi_m\|^2(g_{ik}g_{jl} - g_{il}g_{jk}) + (\psi_i\psi_k g_{jl} - \psi_i\psi_l g_{jk}) \\ &+ (g_{ik}\psi_j\psi_l - g_{il}\psi_j\psi_k)\}, \end{aligned} \tag{5.5}$$

$$R_{ij} - \tilde{R}_{ij} = -\tfrac{1}{9}\{n\|\psi_k\|^2 g_{ij} + (n-2)\psi_i\psi_j\}, \tag{5.6}$$

$$R - \tilde{R} = -\tfrac{1}{9}(n-1)(n+2)\|\psi_i\|^2. \tag{5.7}$$

We next prove the following proposition:

Proposition 5.1. *Suppose* $\alpha = \mathrm{Sym}(\psi \otimes g)$. *Then* $\mathbf{D}\alpha = 0$ *if and only if* $\nabla_o\psi = \frac{1}{2}fg$ *for some* $f \in C^\infty(M)$.

Proof. First we assume $\nabla_0\psi = \frac{1}{2}fg$ for some $f \in C^\infty(M)$, that is to say $\psi_{i,j} = \frac{1}{2}fg_{ij}$ for $f \in C^\infty(M)$. Covariantly differentiating (5.1) and substituting the last expression, we obtain

$$\begin{aligned} \alpha_{ijk,l} &= \tfrac{1}{3}(\psi_{i,l}g_{jk} + \psi_{j,l}g_{ki} + \psi_{k,l}g_{ij}) \\ &= \frac{f}{6}(g_{il}g_{jk} + g_{jl}g_{ki} + g_{kl}g_{ij}), \end{aligned}$$

and we clearly have

$$\alpha_{ijk,l} = \alpha_{ijl,k}.$$

Next we assume $\alpha_{ijk,l} = \alpha_{ijl,k}$. Then from (5.2) we observe that

$$\begin{aligned} 0 &= g^{ij}(\alpha_{ijk,l} - \alpha_{ijl,k}) = \alpha_{k,l} - \alpha_{l,k} \\ &= \frac{n+2}{3}(\psi_{k,l} - \psi_{l,k}), \end{aligned} \tag{5.8}$$

which implies that the 1-form ψ is closed. We now compute

$$\begin{aligned} 0 &= g^{ij}(\alpha_{ijk,l} - \alpha_{ijl,k}) \\ &= \alpha_{i,l} - \tfrac{1}{3}g^{jk}(\psi_{i,k}g_{jl} + \psi_{j,k}g_{li} + \psi_{l,k}g_{ij}) \\ &= \tfrac{1}{3}(n+2)\psi_{i,l} - \tfrac{1}{3}\psi_{i,l} - \tfrac{1}{3}(g^{jk}\psi_{j,k})g_{li} - \tfrac{1}{3}\psi_{l,i} \\ &= \frac{n}{3}\left\{\psi_{i,l} - \frac{1}{n}(g^{jk}\psi_{j,k})g_{il}\right\}, \quad (\psi_{i,l} = \psi_{l,i} \text{ follows from (5.8)}) \end{aligned}$$

and observe that our claim follows by taking $f = (2/n)g^{ij}\psi_{i,j} \in C^\infty(M)$. □

Remark. We have seen that if $\alpha = \mathrm{Sym}(\psi \otimes g)$, then $\mathbf{D}\alpha = 0$ is equivalent to requiring that ψ satisfies the equation

$$\nabla_0 \psi = \frac{f}{2} g \qquad \text{for some } f \in C^\infty(M). \tag{5.9}$$

Note that the metric tensor g canonically identifies $\boldsymbol{\Omega}^1(M)$ with $\mathcal{X}(M)$ as $\psi \mapsto X^\psi \in \mathcal{X}(M)$, $X \mapsto \psi^X \in \boldsymbol{\Omega}^1(M)$, where $\psi \in \boldsymbol{\Omega}^1(M)$, $X \in \mathcal{X}(M)$, and if $\mathcal{L}_X$ denotes the Lie derivative with respect to X, ψ satisfies (5.9) if and only if $d\psi = 0$ and $\mathcal{L}_{X^\psi} g = fg$ for some $f \in C^\infty$. Thus there is a bijective correspondence between the set of solutions to (5.9) and the set of conformal Killing fields X such that $d\psi^X = 0$.

The following theorem in [5] suggests that the existence of a non-isometric conformal Killing field is a fairly strong condition on (M, g).

Theorem 5.2. *Let (M,g) be a connected n-dimensional Riemannian manifold for which $\mathcal{C}^0(M) \neq \mathcal{I}^0(M)$ where $\mathcal{C}^0(M)$ denotes the largest connected group of conformal transformations of (M,g) and $\mathcal{I}^0(M)$ denotes the largest connected group of isometries of (M,g). Then:*

(1) *If M is compact, there is no harmonic p-form of constant length for $1 \leqslant p \leqslant n$ (Goldberg and Kobayashi).*

(2) *If (M,g) is compact and homogeneous, then (M,g) is isometric to a sphere provided $n > 3$ (Goldberg and Kobayashi).*

(3) *If (M,g) is a complete Riemannian manifold of dimension $n \geqslant 3$ with parallel Ricci tensor, then M is isometric to a sphere (Nagano).*

(4) *(M,g) cannot be a compact Riemannian manifold with constant non-positive scalar curvature (Yano and Lichnerowicz).*

Example. Let us give an example of (M,g) admitting non-trivial solutions to the equation (5.9). Consider n-sphere $\mathbf{S}^n$ with its standard metric, and define $\sigma(x) = \langle x, a\rangle \in C^\infty(\mathbf{S}^n)$, where $x \in \mathbf{S}^n \subset \mathbb{R}^{n+1}$, a is a vector in $\mathbb{R}^{n+1}$, and $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product on $\mathbb{R}^{n+1}$. Then, as shown in [5], σ satisfies

$$\nabla_o d\sigma = -\frac{1}{r^2}\sigma g,$$

and also Obata's equation

$$\Delta\sigma = -\frac{n}{r^2}\sigma.$$

Now $\psi = d\sigma$ clearly satisfies (5.9).

A statistical manifold (M, g, ∇) is said to be ∇-*complete* whenever ∇ is geodesically complete. The following example shows that (M, g, ∇) may not be ∇-complete even if (M, g) is complete Riemannian.

Example. Consider $(\mathbb{R}^2, ds^2, \nabla)$, where $ds^2 = dx^2 + dy^2$ and ∇ is determined by the tensor

$$\alpha = \begin{pmatrix} dy & dx \\ dx & -dy \end{pmatrix}.$$

(Incidentally, we have $\operatorname{Tr}\nabla \equiv 0$ and $\tilde{R} = -2$.) We observe that the equations of geodesics $\gamma(t) = (x(t), y(t))$ of ∇ are given by

$$\ddot{x} + 2\dot{x}\dot{y} = 0,$$
$$\ddot{y} + (\dot{x})^2 - (\dot{y})^2 = 0,$$

where " $\dot{}$ " indicates differentiation with respect to the parameter t. We now verify that the pair

$$x(t) = \text{const},$$
$$y(t) = y(t_0) - \ln\{1 - \dot{y}(t_0)(t - t_0)\}$$

determines a geodesic which cannot be extended to all of $\mathbb{R}$. Furthermore, since the whole situation descends to 2-torus $\mathbf{T}^2$ via a covering projection $\mathbb{R}^2 \to \mathbf{T}^2$, it follows that (M, g, ∇) *may not be* ∇*-complete even if* (M, g) *is compact Riemannian.*

6. Special statistical manifolds and ∇-completeness

Let (M, g, ∇) be a statistical manifold and let $\gamma(t)$ be a C^∞ curve in M. We call $\gamma(t)$ a ∇-*geodesic* whenever $\gamma(t)$ is a geodesic of ∇. Observe that the metric g turns the difference tensor $\alpha \in \Omega^1(\operatorname{End} TM)$ into a tri-linear map, which shall also be denoted by α, in a natural manner. We now prove the following proposition relating the time-derivative $\mathrm{d}/\mathrm{d}t\,\|\dot{\gamma}(t)\|^2$ to the tri-linear map α:

Proposition 6.1. *Let* $\gamma(t)$ *be a* ∇*-geodesic on a statistical manifold* (M, g, ∇)*. Then we have*

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\dot{\gamma}(t)\|^2 = -2\alpha(\dot{\gamma}(t), \dot{\gamma}(t), \dot{\gamma}(t)).$$

Proof. Using the notations in Section 2 we have

$$\Gamma^i{}_{jk} = \bar{\Gamma}^i{}_{jk} + \alpha^i{}_{jk} \tag{6.1}$$

in local coordinates. Since $\gamma(t)$ is a ∇-geodesic, it by definition satisfies the equations

$$\ddot{\gamma}^i + \Gamma^i{}_{jk}\dot{\gamma}^j\gamma^k = 0,$$

which are equivalent to the equations

$$\ddot{\gamma}^i + \bar{\Gamma}^i{}_{jk}\dot{\gamma}^j\gamma^k = -\alpha^i{}_{jk}\dot{\gamma}^j\dot{\gamma}^k \tag{6.2}$$

in view of (6.1).

We now compute

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\|\dot\gamma\|^2 &= \frac{\mathrm{d}}{\mathrm{d}t}(g_{ij}\dot\gamma^i\dot\gamma^j)\\ &= \frac{\partial g_{ij}}{\partial x^l}\dot\gamma^l\dot\gamma^i\dot\gamma^j + g_{ij}\ddot\gamma^i\dot\gamma^j + g_{ij}\dot\gamma^i\ddot\gamma^j\\ &= (\bar\Gamma_{ilj} + \bar\Gamma_{jli})\dot\gamma^l\dot\gamma^i\dot\gamma^j + g_{ij}\ddot\gamma^i\dot\gamma^j + g_{ij}\dot\gamma^i\ddot\gamma^j,\end{aligned}$$

where $\bar\Gamma_{ijk} = g_{il}\bar\Gamma^l{}_{jk}$,

$$\begin{aligned}&= \dot\gamma^j(g_{rj}\ddot\gamma^r + \bar\Gamma_{ilj}\dot\gamma^i\dot\gamma^l) + \dot\gamma^i(g_{ir}\ddot\gamma^r + \bar\Gamma_{jli}\dot\gamma^j\dot\gamma^l) \quad \text{by (6.2)}\\ &= -2\alpha_{ijk}\dot\gamma^i\dot\gamma^j\dot\gamma^k. \quad \square\end{aligned}$$

We next prove a lemma concerning the magnitude of $\|\dot\gamma(t)\|$, where $\gamma(t)$ is a ∇-geodesic on a statistical manifold (M, g, ∇).

Lemma 6.2. *Let $\gamma(t)$ be a ∇-geodesic on a statistical manifold (M, g, ∇), and let t_0 be a point in the domain of $\gamma(t)$. Suppose $\alpha = \mathrm{Sym}(d\sigma \otimes g)$ for some $\sigma \in C^\infty(M)$. Then we have*

$$\|\dot\gamma(t)\| = \|\dot\gamma(t_0)\|e^{-\{\sigma(\gamma(t))-\sigma(\gamma(t_0))\}}.$$

Proof. Proposition 6.1 implies

$$\begin{aligned}\frac{d}{dt}\|\dot\gamma\|^2 &= -2\sigma_{,i}g_{jk}\dot\gamma^i\dot\gamma^j\dot\gamma^k = -2(\sigma_{,i}\dot\gamma^i)\|\dot\gamma\|^2\\ &= -2\frac{\mathrm{d}}{\mathrm{d}t}\{\sigma(\gamma(t))\}\|\dot\gamma\|^2,\end{aligned}$$

and now the conclusion of Lemma 6.2 follows. □

We now establish the following theorem about ∇-completeness of a statistical manifold:

Theorem 6.3. *Let (M, g, ∇) be a statistical manifold such that (M, g) is complete Riemannian, $\alpha = \mathrm{Sym}(d\sigma \otimes g)$ for some $\sigma \in C^\infty(M)$, and $L = \inf_M \sigma$ exists. Then (M, g, ∇) is ∇-complete.*

Proof. It suffices to show that every ∇-geodesic $\gamma : [a, b) \to M$ can be extended past $t = b$. Suppose $\gamma(t)$ is defined at $t = t_0$. Then by Lemma 6.2 we obtain

$$\begin{aligned}\|\dot\gamma(t)\| &= \|\dot\gamma(t_0)\|e^{-\{\sigma(\gamma(t))-\sigma(\gamma(t_0))\}}\\ &\leqslant \|\dot\gamma(t_0)\|e^{\{\sigma(\gamma(t_0))-L\}}.\end{aligned} \tag{6.3}$$

Setting $C(\gamma, t_0) = \|\dot\gamma(t_0)\|e^{\{\sigma(\gamma(t_0))-L\}} > 0$, (6.3) can simply be written

$$\|\dot\gamma(t)\| \leqslant C(\gamma, t_0). \tag{6.4}$$

Choose a convergent sequence $t_n \uparrow b$. Then we have from (6.4)

$$d(\gamma(t_m), \gamma(t_n)) \leqslant \int_{t_n}^{t_m} \|\dot{\gamma}(t)\| dt \leqslant C(\gamma, t_0)|t_m - t_n|, \tag{6.5}$$

where $d(\cdot\,,\cdot)$ is the metric determined by g in the usual manner. We note that (6.5) implies that $\{\gamma(t_n)\}$ is a Cauchy sequence with respect to $d(\cdot\,,)$. Since $d(\cdot\,,\cdot)$ is assumed to be complete, we have $\gamma(t_n) \to p \in M$.

Fact (see for example [5]). Let ∇ be an affine connection on (M, g). Then, for every point $p \in M$, there is an open neighborhood W and a positive number $\epsilon > 0$ such that

(1) For any two points q, $q' \in W$, there exists a unique vector $\mathbf{v} \in M_q$ (the fibre of TM through q) such that $\|\mathbf{v}\| < \epsilon$ and $\exp_q \mathbf{v} = q'$, where exp denotes the exponential map of ∇.

(2) For each $q \in W$, the map $\exp_q$ maps the open ϵ-ball in M_q diffeomorphically onto an open set $U_q \supset W$.

For the limit p of the sequence $\{\gamma(t_n)\}$, we choose an open neighborhood W and a positive number $\epsilon > 0$ so that the conclusions (1) and (2) in Fact above hold. Note that we may choose sufficiently large numbers n, m so that the following conditions hold at the same time: $t_n < t_m$; $\gamma(t_n), \gamma(t_m) \in W$;

$$|t_m - t_n| C(\gamma, t_0) < \epsilon. \tag{6.6}$$

We now fix the numbers n, m chosen above and define $\tilde{\gamma}(s) = \gamma(t(s))$, where $t(s) = (t_m - t_n)s + t_n$. Observe that the inverse of $t(s)$ equals

$$s(t) = \frac{t - t_n}{t_m - t_n}$$

so that

$$\tilde{\gamma}(0) = \gamma(t_n) \quad \text{and} \quad \tilde{\gamma}(1) = \gamma(t_m), \tag{6.7}$$

and also observe that $\dot{\tilde{\gamma}}(0) = (t_m - t_n)\dot{\gamma}(t_n)$. The last expression together with (6.6) establishes

$$\|\dot{\tilde{\gamma}}(0)\| = |t_m - t_n| \|\dot{\gamma}(t_n)\| \leqslant |t_m - t_n| C(\gamma, t_0) < \epsilon. \tag{6.8}$$

According to (1) in Fact there exists a unique vector $\mathbf{v} \in M_{\tilde{\gamma}(0)}$ such that $\|\mathbf{v}\| < \epsilon$ and $\exp_{\tilde{\gamma}(0)} \mathbf{v} = \tilde{\gamma}(1)$. By the definition of exponential map, together with (6.7), we obtain

$$\tilde{\gamma}(s) = \exp_{\tilde{\gamma}(0)} s\mathbf{v} \tag{6.9}$$

for $a' \leqslant s < b'$ where $a' = s(a)$, $b' = s(b)$. Setting $s_i = s(t_i)$ we clearly have $s_i \uparrow b'$. We thus only need to show that $\tilde{\gamma}(s)$ can be extended past $s = b'$. By our construction we have $\tilde{\gamma}(s_i) = \gamma(t_i)$ and thus have $\tilde{\gamma}(s_i) \to p$. Let $B_\epsilon(\tilde{\gamma}(0)) \subset M_{\tilde{\gamma}(0)}$ be the ϵ-ball at $\tilde{\gamma}(0)$ such that $\exp_{\tilde{\gamma}(0)} B_\epsilon(\tilde{\gamma}(0)) \supset W \ni p$. Then there exists a unique $\xi \in B_\epsilon(\tilde{\gamma}(0))$ such that $\exp_{\tilde{\gamma}(0)} \xi = p$, and similarly for each i there exists a unique $\xi_i \in B_\epsilon(\tilde{\gamma}(0))$ such that $\exp_{\tilde{\gamma}(0)} \xi_i = \tilde{\gamma}(s_i)$. Now, since $\exp_{\tilde{\gamma}(0)} : B_\epsilon(\tilde{\gamma}(0)) \to U_{\tilde{\gamma}(0)} \supset W$ is a diffeomorphism and $\tilde{\gamma}(s_i) \to p$, we must have $\xi_i \to \xi$, and this means that from (6.9) that $\xi_i = s_i \mathbf{v}$ for all i. We then deduce that $\xi = b'\mathbf{v}$ and $\tilde{\gamma}(s) = \exp_{\tilde{\gamma}(0)} s\mathbf{v}$ clearly extends past $s = b'$. □

The following corollary is a trivial consequence of Theorem 6.3:

Corollary 6.4. *Let (M, g, ∇) be a special statistical manifold with $\alpha = \mathrm{Sym}(\psi \otimes g)$. Suppose $d\psi = 0$ and the universal cover $\tilde{M}$ is compact. Then (M, g, ∇) is ∇-complete.*

We next exhibit some examples of special conjugate symmetric statistical manifolds.

Example 1. $M = \mathbb{R}^2$ with $g_{ij} = \delta_{ij}$. Since all solutions to $\nabla_0\psi = \frac{1}{2}(\mathrm{Tr}\,\nabla_0\psi)g$ are exact, we only need to solve $\sigma_{,ij} - \frac{1}{2}\Delta\sigma\delta_{ij} = 0$. The last equations in this case are just

$$\begin{aligned} &\sigma_{,11} - \sigma_{,22} = 0, \\ &\sigma_{,12} = 0, \end{aligned} \tag{6.10}$$

where $(x^1, x^2) = (x, y)$ are the standard coordinates on $\mathbb{R}^2$. The general solutions to (6.10) are of the form

$$\sigma(x, y) = a(x^2 + y^2) + bx + cy + d,$$

where a, b, c, d are real constants. Note that $\sigma_{,1} = 2ax + b$, $\sigma_{,2} = 2ay + c$, and hence $\alpha_{111} = 2ax + b$, $\alpha_{112} = \alpha_{121} = \alpha_{211} = \frac{1}{3}(2ay + c)$, $\alpha_{122} = \alpha_{212} = \alpha_{221} = \frac{1}{3}(2ax + b)$, $\alpha_{222} = 2ay + c$. We compute

$$\begin{aligned} &\mathrm{Tr}\,\alpha = \tfrac{4}{3}\{(2ax + b)dx + (2ay + c)dy\}, \\ &\|\alpha_i\|^2 = \tfrac{16}{9}\{(2ax + b)^2 + (2ay + c)^2\}, \\ &\|\alpha_{ijk}\|^2 = \tfrac{4}{3}\{(2ax + b)^2 + (2ay + c)^2\}, \\ &\tilde{R} = \tfrac{4}{9}\{(2ax + b)^2 + (2ay + c)^2\}. \end{aligned}$$

We remark that if $a > 0$, then (M, g, ∇) is ∇-complete.

Example 2. $M = \mathbf{H}^2$ with Poincaré metric

$$g_{ij} = \frac{1}{y^2}\delta_{ij}.$$

We must solve $\sigma_{,ij} - \frac{1}{2}\Delta\sigma\delta_{ij} = 0$ or

$$\sigma_{,11} - \sigma_{,22} = 0, \quad \sigma_{,12} = 0,$$

which can also be written as

$$\begin{aligned} &\frac{\partial^2\sigma}{\partial x^2} - \frac{\partial^2\sigma}{\partial y^2} = \frac{2}{y}\frac{\partial\sigma}{\partial y}, \\ &\frac{\partial^2\sigma}{\partial x\partial y} + \frac{1}{y}\frac{\partial\sigma}{\partial x} = 0. \end{aligned} \tag{6.11}$$

The general solutions to (6.11) are given by

$$\sigma(x,y)=\frac{ax+b}{y}+c,$$

where a,b,c are real constants. Observing $\sigma_{,1}=a/y$, $\sigma_{,2}=-(ax+b)/y^2$, we obtain $\alpha_{111}=a/y^3$, $\alpha_{112}=\alpha_{121}=\alpha_{211}=-\frac{1}{3}(ax+b)/y^4$, $\alpha_{122}=\alpha_{212}=\alpha_{221}=\frac{1}{3}a/y^3$, $\alpha_{222}=-(ax+b)/y^4$. We compute

$$\begin{aligned}
\operatorname{Tr}\alpha &= \frac{4a}{3y}dx-\frac{4(ax+b)}{3y^2}dy,\\
\|\alpha_i\|^2 &= \frac{16}{9}\frac{a^2(x^2+y^2)+2abx+b^2}{y^2},\\
\|\alpha_{ijk}\|^2 &= \frac{4}{3}\frac{a^2(x^2+y^2)+2abx+b^2}{y^2},\\
\tilde{R} &= \frac{4}{9}\frac{a^2(x^2+y^2)+2abx+b^2}{y^2}-2.
\end{aligned}$$

Note that if $a=0$, $b>0$, then (M,g,∇) is ∇-complete.

Example 3. $M=\mathbf{S}^2$ with g the canonical metric on $\mathbf{S}^2$. We use local coordinates (x,y) such that the metric takes the form

$$g=\frac{4(dx^2+dy^2)}{(1+x^2+y^2)^2}.$$

We must solve $\sigma_{,ij}-\frac{1}{2}\Delta\sigma\delta_{ij}=0$ or

$$\sigma_{,11}-\sigma_{,22}=0,\qquad \sigma_{,12}=0,$$

which are equivalent to

$$\begin{aligned}
&\frac{\partial^2\sigma}{\partial x^2}-\frac{\partial^2\sigma}{\partial y^2}+\frac{4}{1+x^2+y^2}\Big(x\frac{\partial\sigma}{\partial x}-y\frac{\partial\sigma}{\partial y}\Big)=0,\\
&\frac{\partial^2\sigma}{\partial x\partial y}+\frac{2y}{1+x^2+y^2}\frac{\partial\sigma}{\partial x}+\frac{2x}{1+x^2+y^2}\frac{\partial\sigma}{\partial y}=0.
\end{aligned}\tag{6.12}$$

If we use a complex coordinate $z=x+\sqrt{-1}y$, (6.12) takes a simple form

$$\frac{\partial^2\sigma}{\partial z^2}+\frac{2\bar{z}}{1+|z|^2}\frac{\partial\sigma}{\partial z}=0.\tag{6.13}$$

Set $H=(1+|z|^2)^2\,\partial\sigma/\partial z$. Then we compute

$$\frac{\partial H}{\partial z}=(1+|z|^2)^2\Big(\frac{\partial^2\sigma}{\partial z^2}+\frac{2\bar{z}}{1+|z|^2}\frac{\partial\sigma}{\partial z}\Big),$$

and we see that σ satisfies (6.13) if and only if $\partial H/\partial z = 0$ or H is anti-holomorphic. Thus we must solve

$$\frac{\partial \sigma}{\partial z} = \frac{H}{(1+|z|^2)^2}, \tag{6.14}$$

where H is an arbitrary anti-holomorphic function. We next set $\sigma = F/(1+|z|^2)$ and compute

$$\frac{\partial \sigma}{\partial z} = \frac{\frac{\partial F}{\partial z}(1+|z|^2) - F\bar{z}}{(1+|z|^2)^2}. \tag{6.15}$$

Comparing (6.14) with (6.15) we deduce that σ satisfies (6.14) if and only if F satisfies

$$H = \frac{\partial F}{\partial z}(1+|z|^2) - F\bar{z}$$

or

$$0 = \frac{\partial H}{\partial z} = \frac{\partial}{\partial z}\left\{\frac{\partial F}{\partial z}(1+|z|^2) - F\bar{z}\right\} = (1+|z|^2)\frac{\partial^2 F}{\partial z^2},$$

and if and only if F satisfies

$$\frac{\partial^2 F}{\partial z^2} = 0. \tag{6.16}$$

Now, since

$$\frac{\partial^2 F}{\partial z^2} = \frac{1}{4}\left\{\left(\frac{\partial^2 F}{\partial x^2} - \frac{\partial^2 F}{\partial y^2}\right) - 2\sqrt{-1}\frac{\partial^2 F}{\partial x \partial y}\right\}$$

and F is real, F satisfies (6.16) if and only if $F = a(x^2+y^2) + bx + cy + d$, where a, b, c, d are real constants.

Thus the general solutions to (6.12) are of the form

$$\sigma = \frac{a(x^2+y^2) + bx + cy + d}{1+x^2+y^2}. \tag{6.17}$$

Let (x^1, x^2, x^3) be the standard coordinates of $\mathbb{R}^3$. Then the so-called stereographic chart is defined by

$$x^1 = \frac{4u}{4+u^2+v^2}, \quad x^2 = \frac{4v}{4+u^2+v^2}, \quad x^3 = \frac{2(u^2+v^2)}{4+u^2+v^2},$$

and our coordinates (x, y) are related to (u, v) by $x = \frac{1}{2}u$, $y = \frac{1}{2}v$. We now recall from Example after Theorem 5.2 that functions $\sigma = x^i = \langle x, e_i \rangle$, with (e_1, e_2, e_3) the natural basis of $\mathbb{R}^3$, considered as functions on $S^2 \subset \mathbb{R}^3$ are all solutions to $\sigma_{,ij} - \frac{1}{2}\Delta\sigma g_{ij} = 0$; in fact,

$$x^1 = \frac{2x}{1+x^2+y^2}, \quad x^2 = \frac{2y}{1+x^2+y^2}, \quad x^3 = \frac{2(x^2+y^2)}{1+x^2+y^2}$$

are just special cases of (6.17).

It is interesting to note the following well-known fact about conformal Killing fields on $\mathbf{S}^n$: all conformal Killing fields on the standard n-sphere $(\mathbf{S}^n, g)$ are obtained by first restricting the Killing fields on $\mathbb{R}^{n+1}$ to $\mathbf{S}^n$ and next projecting them onto $T\mathbf{S}^n$. Four independent solutions in (6.17) indeed correspond to the translations along the coordinate axes of $\mathbb{R}^3$.

Needless to say, $(\mathbf{S}^n, g, \nabla)$ with $\sigma = \langle x, a\rangle$, $a \in \mathbb{R}^{n+1}$, will be a ∇-complete special conjugate symmetric statistical manifold.

Finally we compute the following quantities for $(\mathbf{S}^2, g, \nabla)$ with σ given by (6.17):

$$\sigma_{,1} = \frac{b(y^2-x^2) - 2cxy + 2(a-d)x + b}{(1+x^2+y^2)^2},$$

$$\sigma_{,2} = \frac{b(x^2-y^2) - 2cxy + 2(a-d)y + b}{(1+x^2+y^2)^2},$$

$$\alpha_{111} = \sigma_{,1}\frac{4}{(1+x^2+y^2)^2},$$

$$\alpha_{112} = \alpha_{121} = \alpha_{211} = \frac{\sigma_{,2}}{3}\frac{4}{(1+x^2+y^2)^2},$$

$$\alpha_{122} = \alpha_{212} = \alpha_{221} = \frac{\sigma_{,1}}{3}\frac{4}{(1+x^2+y^2)^2},$$

$$\alpha_{222} = \sigma_{,2}\frac{4}{(1+x^2+y^2)^2},$$

$$\operatorname{Tr}\alpha = \tfrac{4}{3}\sigma_{,1}dx + \tfrac{4}{3}\sigma_{,2}dy,$$

$$\|\alpha_i\|^2 = \tfrac{4}{9}(\sigma_{,1}^2 + \sigma_{,2}^2)(1+x^2+y^2)^2,$$

$$\|\alpha_{ijk}\|^2 = \tfrac{1}{3}(\sigma_{,1}^2 + \sigma_{,2}^2)(1+x^2+y^2)^2,$$

$$\tilde{R} = -\tfrac{1}{9}(\sigma_{,1}^2 + \sigma_{,2}^2)(1+x^2+y^2)^2.$$

References

[1] S. Amari, *Differential-Geometrical Methods in Statistics*, Lecture Notes in Statistics **28** (Springer, New York, 1985).

[2] C. Atkinson and A.F.S. Mitchell, Rao's distance measure, *Sankhya A* **43** (1981) 345–365.

[3] E. Calabi, Improper affine hypersurface of convex type and a generalization of a theorem by K. Jörgens, *Michigan Math. J.* **5** (1958) 105–126.

[4] S.Y. Cheng and S.T. Yau, Differential equations on Riemannian manifolds and their geometric applications, *Comm. Pure Appl. Math.* **28** (1975) 333–354.

[5] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry I-II* (Wiley, New York, 1963).

[6] T. Kurose, Dual connections and affine geometry, *Mathematische Zeitschrift* **203** (1990) 115–121.

[7] S.L. Lauritzen, Statistical manifolds, *IMS Lecture Notes-Monograph Series* **10** (1987) 163–216.

[8] H. Shima, Compact locally Hessian manifolds, *Osaka J. Math.* **15** (1978) 509–513.

[9] L.T. Skovgaard, A Riemannian geometry of the multivariate normal model, *Scand. J. Statist.* **11** (1984) 211–223.