HIROSHIMA MATH. J.
30 (2000), 403–414

Foliations and divergences of flat statistical manifolds

Keiko UOHASHI, Atsumi OHARA and Takao FUJII

(Received July 30, 1998)
(Revised November 10, 1999)

ABSTRACT. A Hessian domain $(\Omega, \tilde{D}, \tilde{g} = \tilde{D}d\varphi)$ is a flat statistical manifold, and level surfaces of φ are 1-conformally flat statistical submanifolds of $(\Omega, \tilde{D}, \tilde{g})$. In this paper we consider a foliation defined by level surfaces of φ and its orthogonal foliation, and then we investigate divergences restricted to leaves of these foliations.

1. Introduction

Statistical manifolds have been studied in terms of information geometry. *Dualistic structures* of statistical manifolds play important roles on statistical inference, control systems theory, and so on [1] [12]. It is known that a *Hessian structure* is a *dually flat structure* and gives, for examples geometry of an exponential family [14]. Applications of the dually flat structures of submanifolds are in [4] [12]. Non-flat statistical manifolds are studied in [6] [7] [8]. It seems that there are not results on *statistical submanifolds* without dually flat structures. So, we treat non-flat dualistic structures on submanifolds, especially on *level surfaces* of *Hessian domain*, and show 1-conformal flatness, if considering a Hessian domain as a flat statistical manifold.

Let φ be a function on a domain Ω in a real affine space $\mathbf{A}^{n+1}$. Denoting by $\tilde{D}$ the canonical flat affine connection on $\mathbf{A}^{n+1}$, we set $\tilde{g} = \tilde{D}d\varphi$ and suppose that $\tilde{g}$ is non-degenerate. Then a Hessian domain $(\Omega, \tilde{D}, \tilde{g})$ is a *flat statistical manifold.* In [15] we proved that n-dimensional level surfaces of φ are 1*-conformally flat statistical submanifolds* of $(\Omega, \tilde{D}, \tilde{g})$. Using this fact, we show that *dual-projectively equivalent* affine connections can be led on a leaf of a foliation $\mathscr{F}$ defined by n-dimensional level surfaces of φ on Ω. In addition we study the *orthogonal foliation* $\mathscr{F}^{\perp}$ of $\mathscr{F}$.

We also discuss *divergences* on leaves of the foliations $\mathscr{F}$ and $\mathscr{F}^{\perp}$ in §4. Nagaoka and Amari first studied divergences of flat statistical manifolds in view of statistics [1]. Kurose defined the canonical divergences of 1-conformally flat statistical manifolds [7]. In this paper we show that, for $M \in \mathscr{F}$, Kurose's

---

2000 *Mathematics Subject Classification.* 53A15

Key words and phrases. Flat statistical manifold, 1-conformally flat statistical submanifold, Hessian domain, level surface, divergence.

divergence of a 1-conformally flat statistical submanifold (M, D, g) coincides with the restriction of Nagaoka and Amari's divergence of $(\Omega, \tilde{D}, \tilde{g})$. In §5, we give the decomposition of the divergence of $(\Omega, \tilde{D}, \tilde{g})$ with respect to orthogonal foliations $\mathscr{F}$ and $\mathscr{F}^{\perp}$, and see that the projection of a point in Ω to M along a leaf of $\mathscr{F}^{\perp}$ is given by minimization of the divergence. Finally we give a *gradient system* using the divergence. Gradient systems are important to study relation with information geometry and integrable dynamical systems [4] [10].

An original reason for our investigation of divergences is that divergences are canonical *contrast functions*, which generate statistical manifolds. On contrast functions and *minimum contrast leaves*, see [2] [9]. Divergences of *conformally-projectively flat statistical manifolds* are described in [8]. Shima studied the Riemannian foliations on Hessian domains deeply in [13].

2. Statistical manifolds and Hessian domains

We recall properties of statistical manifolds, Hessian domains, and affine differential geometry.

Let $\tilde{D}$ and $\{x^1, \dots, x^{n+1}\}$ be the canonical flat affine connection and the canonical affine coordinate system on $\mathbf{A}^{n+1}$, i.e., $\tilde{D}dx^i = 0$. If the Hessian $\tilde{D}d\varphi = \sum_{i,j}(\partial^2\varphi/\partial x^i\partial x^j)dx^i dx^j$ is non-degenerate for a function φ on a domain Ω in $\mathbf{A}^{n+1}$, we call $(\Omega, \tilde{D}, \tilde{g} = \tilde{D}d\varphi)$ a Hessian domain. For a torsion-free affine connection ∇ and a pseudo-Riemannian metric h on a manifold N, the triple (N, ∇, h) is called a statistical manifold if ∇h is symmetric. If the curvature tensor R of ∇ vanishes, (N, ∇, h) is said to be flat. A Hessian domain $(\Omega, \tilde{D}, \tilde{g} = \tilde{D}d\varphi)$ is a flat statistical manifold. Conversely, a flat statistical manifold is locally a Hessian domain [1] [13].

For a statistical manifold (N, ∇, h), let ∇' be an affine connection on N such that

$$Xh(Y, Z) = h(\nabla_X Y, Z) + h(Y, \nabla'_X Z), \qquad X, Y, Z \in \mathscr{X}(N),$$

where $\mathscr{X}(N)$ is the set of all tangent vector fields on N. The affine connection ∇' is torsion free, and $\nabla' h$ symmetric. Then ∇' is called the dual connection of ∇, the triple (N, ∇', h) the dual statistical manifold of (N, ∇, h), and (∇, ∇', h) the dualistic structure on N, respectively. The curvature tensor of ∇' vanishes if and only if one of ∇ does, and then (∇, ∇', h) is called the dually flat structure.

Let $\mathbf{A}^*_{n+1}$ and $\{x^*_1, \dots, x^*_{n+1}\}$ be the dual affine space of $\mathbf{A}^{n+1}$ and the dual affine coordinate system of $\{x^1, \dots, x^{n+1}\}$, respectively. We define the *gradient mapping* $\tilde{\iota}$ from Ω to $\mathbf{A}^*_{n+1}$ by

$$x_i^* \circ \tilde{\imath} = -\frac{\partial \varphi}{\partial x^i}, \tag{1}$$

and a flat affine connection $\tilde{D}'$ on Ω by

$$\tilde{\imath}_*(\tilde{D}'_{\tilde{X}}\tilde{Y}) = \tilde{D}^*_{\tilde{X}}\tilde{\imath}_*(\tilde{Y}) \qquad \text{for } \tilde{X}, \tilde{Y} \in \mathscr{X}(\Omega),$$

where $\tilde{D}^*_{\tilde{X}}\tilde{\imath}_*(\tilde{Y})$ is covariant derivative along $\tilde{\imath}$ induced by the canonical flat affine connection $\tilde{D}^*$ on $\mathbf{A}^*_{n+1}$. Then $(\Omega, \tilde{D}', \tilde{g})$ is the dual statistical manifold of $(\Omega, \tilde{D}, \tilde{g})$. We set $x'_i = x^*_i \circ \tilde{\imath} = -(\varphi/x^i)$. Then $\{x'_1, \dots, x'_{n+1}\}$ is the affine flat coordinate system with respect to $\tilde{D}'$, i.e., $\tilde{D}'dx'_i = 0$. Remark that a straight line with respect to an affine coordinate $\{x^1, \dots, x^{n+1}\}$ (resp. $\{x'_1, \dots, x'_{n+1}\}$) is a $\tilde{D}$- (resp. $\tilde{D}'$-) geodesic, where we call a geodesic relative to $\tilde{D}$ (resp. $\tilde{D}'$) a $\tilde{D}$- (resp. $\tilde{D}'$-) geodesic. If $\tilde{\imath}$ is invertible, we can define a function on $\Omega^* = \tilde{\imath}(\Omega)$ called the *Legendre transform* φ^* of φ by

$$\varphi^* \circ \tilde{\imath} = -\sum_i x^i x'_i - \varphi.$$

For $\alpha \in \mathbf{R}$, statistical manifolds (N, ∇, h) and $(N, \overline{\nabla}, \bar{h})$ are said to be α-conformally equivalent if there exists a function ϕ on N such that

$$\bar{h}(X, Y) = e^{\phi} h(X, Y),$$

$$h(\overline{\nabla}_X Y, Z) = h(\nabla_X Y, Z) - \frac{1+\alpha}{2} d\phi(Z) h(X, Y) + \frac{1-\alpha}{2}\{d\phi(X)h(Y, Z) + d\phi(Y)h(X, Z)\}$$

for $X, Y, Z \in \mathscr{X}(N)$. A statistical manifold (N, ∇, h) is called α-conformally flat if (N, ∇, h) is locally α-conformally equivalent to a flat statistical manifold. Statistical manifolds (N, ∇, h) and $(N, \overline{\nabla}, \bar{h})$ are α-conformally equivalent if and only if the dual statistical manifolds (N, ∇', h) and $(N, \overline{\nabla}', \bar{h})$ are $(-\alpha)$-conformally equivalent. Especially, a statistical manifold (N, ∇, h) is 1-conformally flat if and only if the dual statistical manifold (N, ∇', h) is (-1)-conformally flat [7].

Henceforth, we suppose that $\tilde{g}$ is positive definite.

Let $\tilde{E}$ be the gradient vector field of φ on Ω defined by

$$\tilde{g}(\tilde{X}, \tilde{E}) = d\varphi(\tilde{X}) \qquad \text{for } \tilde{X} \in \mathscr{X}(\Omega),$$

where $\mathscr{X}(\Omega)$ is the set of all tangent vector fields on Ω. We set

$$\Omega_o = \{p \in \Omega \mid d\varphi_p \neq 0\},$$

$$E = -d\varphi(\tilde{E})^{-1}\tilde{E} \qquad \text{on } \Omega_o.$$

For $p \in \Omega_o$, E_p is perpendicular to T_pM with respect to $\tilde{g}$, where $M \subset \Omega_o$ is a level surface of φ containing p and T_pM is the set of all tangent vectors at p on M.

Let x be a canonical immersion of an n-dimensional level surface M into Ω. For $\tilde{D}$ and an affine immersion (x, E), we can define the induced affine connection D^E, the affine fundamental form g^E on M by

$$\tilde{D}_X Y = D^E_X Y + g^E(X, Y)E \qquad \text{for } X, Y \in \mathscr{X}(M).$$

For M, we denote by (M, D, g) the statistical submanifold realized in $(\Omega, \tilde{D}, \tilde{g})$, which coincides with the manifold (M, D^E, g^E) induced by an affine immersion (x, E) [15].

An affine immersion (x, E) is non-degenerate equiaffine, and (M, D, g) is a 1-conformally flat statistical manifold [7]. In fact, we have:

THEOREM 2.1. ([15]) *Let M be a simply connected n-dimensional level surface of φ on an $(n+1)$-dimensional Hessian domain $(\Omega, \tilde{D}, \tilde{g} = \tilde{D}d\varphi)$ with a Riemannian metric $\tilde{g}$, and suppose that $n \geq 2$. If we consider $(\Omega, \tilde{D}, \tilde{g})$ a flat statistical manifold, (M, D, g) is a 1-conformally flat statistical submanifold of $(\Omega, \tilde{D}, \tilde{g})$, where we denote by D and g the connection and the Riemannian metric on M induced by $\tilde{D}$ and $\tilde{g}$.*

3. Foliations by level surfaces

We denote by $\mathscr{F}$ and $\mathscr{F}^{\perp}$ a foliation on Ω_o defined by level surfaces of φ and a foliation by the flow of $\tilde{E}$, respectively. In this section we relate these orthogonal foliations with the dualistic structure $(\tilde{D}, \tilde{D}', \tilde{g})$.

Let $M, \hat{M}$ be two leaves of $\mathscr{F}$, and $(M, D, g), (\hat{M}, \hat{D}, \hat{g})$ two statistical submanifolds of $(\Omega, \tilde{D}, \tilde{g})$. We denote by E the vector field on Ω_o defined in §2, and by $\iota, \hat{\iota}$ the restriction of $\tilde{\iota}$ to $M, \hat{M}$, respectively. Non-degenerate affine immersions $(x, E), (\hat{x}, E)$ realize $(M, D, g), (\hat{M}, \hat{D}, \hat{g})$ in $\mathbf{A}^{n+1}$, where $x, \hat{x}$ are canonical immersions of $M, \hat{M}$ into Ω, respectively.

Then ι is the conormal immersion for x. In fact, denoting by $\langle a, b \rangle$ a pairing of $a \in \mathbf{A}^*_{n+1}$ and $b \in \mathbf{A}^{n+1}$, we have

$$\langle \iota(p), Y_p \rangle = 0 \qquad \text{for } Y_p \in T_pM, \qquad \langle \iota(p), E_p \rangle = 1$$

for $p \in M$, considering $T_p\mathbf{A}^{n+1}$ with $\mathbf{A}^{n+1}$. Moreover, ι satisfies that

$$\langle \iota_*(Y), E \rangle = 0, \qquad \langle \iota_*(Y), X \rangle = -g(Y, X)$$

and

$$\tilde{D}^*_X \iota_*(Y) = \iota_*(D'_X Y) - g'(X, Y)\iota$$

for $X, Y \in \mathscr{X}(M)$, where D' is the dual connection of D and g' the second fundamental form. Since g is non-degenerate, an immersion $\iota : M \to \mathbf{A}^*_{n+1} - \{0\}$ is a centro-affine hypersurface. Similarly a conormal immersion $\hat{\iota} : \hat{M} \to \mathbf{A}^*_{n+1} - \{0\}$ for $\hat{x}$ is also a centro-affine hypersurface [11].

We set $(e^\lambda)(p) = e^{\lambda(p)}$ for $p \in M$ and the function λ on M such that $e^{\lambda(p)}\iota(p) \in \hat{\iota}(\hat{M})$. We define a mapping $\pi : M \to \hat{M}$ by

$$\hat{\iota} \circ \pi = e^\lambda \iota.$$

We denote by $\bar{D}'$ an affine connection on M defined by

$$\pi_*(\bar{D}'_X Y) = \hat{D}'_{\pi_*(X)}\pi_*(Y) \qquad \text{for } X, Y \in \mathscr{X}(M),$$

and by $\bar{g}$ a Riemannian metric on M such that

$$\bar{g}(X, Y) = e^\lambda g(X, Y). \tag{2}$$

THEOREM 3.1. *For affine connections $D', \bar{D}'$ on M, we have*
(i) *D' and $\bar{D}'$ are projectively equivalent.*
(ii) *(M, D', g) and $(M, \bar{D}', \bar{g})$ are (-1)-conformally equivalent.*

PROOF. By definition of π, $\bar{D}'$ is the connection on M induced by $e^\lambda \iota$. Since D' is induced by ι, from a property of centro-affine hypersurfaces, it follows (cf. [11]) that

$$\bar{D}'_X Y = D'_X Y + d\lambda(X)Y + d\lambda(Y)X. \tag{3}$$

Thus (i) holds.

Statistical manifolds (M, D', g) and $(M, \bar{D}', \bar{g})$ are by definition (-1)-conformally equivalent if they satisfy (2) and (3). Thus (ii) holds. □

We denote by $\bar{D}$ an affine connection on M defined by

$$\pi_*(\bar{D}_X Y) = \hat{D}_{\pi_*(X)}\pi_*(Y) \qquad \text{for } X, Y \in \mathscr{X}(M).$$

From duality of $\hat{D}$ and $\hat{D}'$, $\bar{D}$ is the dual connection of $\bar{D}'$ on M. Then the next theorem holds.

THEOREM 3.2. *For affine connections $D, \bar{D}$ on M, we have*
(i) *D and $\bar{D}$ are dual-projectively equivalent.*
(ii) *(M, D, g) and $(M, \bar{D}, \bar{g})$ are 1-conformally equivalent.*

PROOF. We have

$$g(\bar{D}_X Y, Z) = g(D_X Y, Z) - d\lambda(Z)g(X, Y) \tag{4}$$

which is equivalent to that (3) holds [7]. Affine connections D and $\bar{D}$ are by definition dual-projectively equivalent if $g(\bar{D}_X Y, Z) = g(D_X Y, Z) - \kappa(Z)g(X, Y)$ for some 1-form κ [5]. Thus (i) holds.

Statistical manifolds (M, D, g) and $(M, \bar{D}, \bar{g})$ are 1-conformally equivalent if they satisfy (2) and (4). Thus (ii) holds. □

For $\mathscr{F}^{\perp}$, we have:

PROPOSITION 3.3. *Every leaf of the foliation $\mathscr{F}^{\perp}$ in the introduction is a $\tilde{D}'$-geodesic on Ω_o under a certain parametrization.*

PROOF. It suffices to see that any integral curve of $\tilde{E}$ is a $\tilde{D}'$-geodesic. To see it, we consider the flow defined by

$$\frac{dx^i}{dt} = \tilde{E}^i \qquad (i = 1, \dots, n+1), \tag{5}$$

where $\tilde{E}^1, \dots, \tilde{E}^{n+1}$ are functions on Ω such that $\tilde{E} = \tilde{E}^i(\partial/\partial x^i)$. By definition we have $\tilde{E}^i = \tilde{g}^{ij}(\partial\varphi/\partial x^j)$, where $\tilde{g}_{ij} = \tilde{g}(\partial/\partial x^i, \partial/\partial x^j)$ and $(\tilde{g}^{ij})$ is the inverse matrix of $(\tilde{g}_{ij})$. Since $x_i' = -\partial\varphi/\partial x^i$, we have $dx^i/dt = -\tilde{g}^{ij}x_j'$, i.e., $-\tilde{g}_{ij}(dx^j/dt) = x_i'$. Moreover $\tilde{g}_{ij} = \partial^2\varphi/\partial x^i\partial x^j$ implies $\tilde{g}_{ij} = -\partial x_i'/\partial x^j$ and

$$\frac{dx_i'}{dt} = x_i' \qquad (i = 1, \dots, n+1).$$

Thus, for an initial point $x'(0) = \{x_1'(0), \dots, x_{n+1}'(0)\} \in \Omega_o$, the integral curve of the flow (5) is described by

$$x_i'(t) = e^t x_i'(0).$$

Hence the integral curve of $\tilde{E}$ is a straight line with respect to an affine coordinate $\{x_1', \dots, x_{n+1}'\}$, and the image of the integral curve is a $\tilde{D}'$-geodesic on Ω_o under a certain parametrization. □

In [1] orthogonal foliations are constructed only by flat submanifolds, and we extended to the case of 1-conformally flat statistical submanifolds. From the proof of Proposition 3.3, we can obtain a leaf of $\mathscr{F}^{\perp}$ by a dilation of a position vector of a point in $\Omega^* = \tilde{\imath}(\Omega)$.

COROLLARY 3.4. *For $p \in L \in \mathscr{F}^{\perp}$ we have*

$$\tilde{\imath}(L) = \{e^t\tilde{\imath}(p) \mid t \in \mathbf{R}\} \cap \Omega^*.$$

4. Divergences and orthogonal foliations

First we define divergences of statistical manifolds.

DEFINITION 4.1. ([1]) The *divergence* ρ of a flat statistical manifold $(\Omega, \tilde{D}, \tilde{g})$ is defined by

$$\rho(p,q) = \varphi(p) + \varphi^*(\tilde{\imath}(q)) + \sum_{i=1}^{n+1} x^i(p)x_i'(q) \qquad \text{for } p,q \in \Omega,$$

where φ^* is the Legendre transform of φ.

DEFINITION 4.2. ([7]) Let (N, ∇, h) be a 1-conformally flat statistical manifold realized by a non-degenerate affine immersion (v, ξ) into $\mathbf{A}^{n+1}$, and w the conormal immersion for v. Then the divergence ρ_{conf} of (N, ∇, h) is defined by

$$\rho_{\mathrm{conf}}(p,q) = \langle w(q), v(p) - v(q) \rangle \qquad \text{for } p,q \in N.$$

The definition of ρ_{conf} is independent of the choice of a realization of (N, ∇, h).

It is known that an arbitrary statistical manifold is induced by a contrast function [9]. These divergences are contrast functions of a flat statistical manifold and of a 1-conformally flat statistical manifold.

For $M \in \mathscr{F}$, we denote by ρ_{conf} the divergence of (M, D, g) induced by a non-degenerate equiaffine immersion (x, E) by Definition 4.2. Since (M, D, g) is a submanifold of $(\Omega, \tilde{D}, \tilde{g} = \tilde{D}d\varphi)$, we can define the divergence ρ_{sub} of (M, D, g) by the restriction of the divergence of $(\Omega, \tilde{D}, \tilde{g})$ defined by Definition 4.1, i.e., $\rho_{\mathrm{sub}}(p,q) = \rho(p,q)$. Then we obtain:

THEOREM 4.3. *For a 1-conformally flat statistical submanifold (M, D, g) of $(\Omega, \tilde{D}, \tilde{g})$, two divergences ρ_{conf} and ρ_{sub} coincide each other.*

For $p \in \Omega$ and $q \in M$, we set $\tilde{\rho}(p,q) = \langle \imath(q), x(p) - x(q) \rangle$, where $\imath$ is the conormal immersion for x. The function $\tilde{\rho}(p, \cdot)$ is called the affine distance function for (x, E) from p. For the proof of Theorem 4.3, we describe the divergence ρ by the affine distance function $\tilde{\rho}$.

LEMMA 4.4. *we have*

$$\rho(p,q) = \varphi(p) - \varphi(q) + \tilde{\rho}(p,q) \qquad \textit{for } p \in \Omega, q \in M.$$

PROOF. Since $\varphi^*(\imath(q)) = -\sum_{i=1}^{n+1} x^i(q)x_i'(q) - \varphi(q)$, it follows that

$$\rho(p,q) = \varphi(p) - \varphi(q) + \sum_{i=1}^{n+1} x_i'(q)(x^i(p) - x^i(q)). \tag{6}$$

Equations $\sum_{i=1}^{n+1} x_i'(q)(x^i(p) - x^i(q)) = \langle \imath(q), x(p) - x(q) \rangle = \tilde{\rho}(p,q)$ imply Lemma 4.4. □

PROOF of THEOREM 4.3. For $p,q \in M$, $\varphi(p) = \varphi(q)$ holds. Since $\rho_{\mathrm{sub}}(p,q) = \rho(p,q)$ and $\rho_{\mathrm{conf}}(p,q) = \tilde{\rho}(p,q)$, by Lemma 4.4 we have

$$\rho_{\mathrm{sub}}(p,q) = \rho_{\mathrm{conf}}(p,q).$$

□

Let us denote both ρ_{sub} and ρ_{conf} by the same notation ρ.

We can apply Lemma 4.4 to a point $q \in \Omega_o$. For a point $r \in \Omega$ such that $d\varphi_r = 0$, $x_i'(r) = 0$ holds, and thus we have $\varphi^*(\iota(r)) = -\varphi(r)$ by the definition of the Legendre transform. Hence we have by Definition 4.1:

COROLLARY 4.5. *For $p \in \Omega$ and $r \in \Omega$ such that $d\varphi_r = 0$, we have*

$$\rho(p,r) = \varphi(p) - \varphi(r).$$

5. Projection by the minimum divergence

We shall describe the decomposition of the divergence of a flat statistical manifold $(\Omega, \tilde{D}, \tilde{g})$ with respect to orthogonal foliations $\mathscr{F}$ and $\mathscr{F}^{\perp}$.

THEOREM 5.1. *Let (M, D, g) be a 1-conformally flat statistical submanifold of an $(n+1)$-dimensional Hessian domain $(\Omega, \tilde{D}, \tilde{g} = \tilde{D}d\varphi)$, where M is an n-dimensional level surface of φ, and let $p, q \in M$, $r \in \Omega$. If a tangent vector at q, of the $\tilde{D}'$-geodesic through q and r, is perpendicular to T_qM with respect to $\tilde{g}$, we have*

$$\rho(p,r) = \mu\rho(p,q) + \rho(q,r), \tag{7}$$

where $\tilde{\iota}$ is the gradient mapping of φ defined by (1) *in* §2 *and $\tilde{\iota}(r) = \mu\tilde{\iota}(q)$, $\mu \in \mathbf{R}$.*

PROOF. Recall that ι is the restriction of $\tilde{\iota}$ to M, and using $x_i' = x_i^* \circ \tilde{\iota}$ and Definition 4.1, we have

$$\begin{aligned}
&\rho(p,q) + \rho(q,r) \\
&\quad = \varphi(p) - \varphi(r) + \sum_{i=1}^{n+1}(x_i'(r) - x_i'(q))(x^i(q) - x^i(p)) + \sum_{i=1}^{n+1} x_i'(r)(x^i(p) - x^i(r)) \\
&\quad = \varphi(p) - \varphi(r) + \langle \tilde{\iota}(r) - \tilde{\iota}(q), x(q) - x(p)\rangle + \langle \tilde{\iota}(r), x(p) - x(r)\rangle.
\end{aligned}$$

By Lemma 4.4, $\rho(p,r) = \varphi(p) - \varphi(r) + \langle \tilde{\iota}(r), x(p) - x(r)\rangle$ holds. Thus we get

$$\rho(p,r) = \rho(p,q) + \rho(q,r) + \langle \tilde{\iota}(r) - \tilde{\iota}(q), x(p) - x(q)\rangle.$$

From Corollary 3.4 the trajectory $C \in \Omega$ of the $\tilde{D}'$-geodesic, through q, r and perpendicular to T_qM, satisfies that

$$\{e^t\tilde{\iota}(q) \mid t \in \mathbf{R}\} \cap \Omega^* \subset \tilde{\iota}(C).$$

Thus there exists a real number μ such that $\tilde{\iota}(r) = \mu\tilde{\iota}(q)$. Since $\rho(p,q) = \langle \tilde{\iota}(q), x(p) - x(q)\rangle$, we have

$$\langle \tilde{\iota}(r) - \tilde{\iota}(q), x(p) - x(q)\rangle = (\mu - 1)\rho(p,q).$$

Thus, we obtain Theorem 5.1. □

By this decomposition we can obtain the projection of a point in Ω_o to $M \in \mathscr{F}$ along a leaf of $\mathscr{F}^{\perp}$.

COROLLARY 5.2. *Let M be an arbitrary leaf of $\mathscr{F}$ and r in $\Omega_o = \{p \in \Omega \mid d\varphi_p \neq 0\}$. Then the unique minimizer of a function $\rho(\cdot, r)$ on M is the intersection point of L_r and M, where L_r is the leaf of $\mathscr{F}^{\perp}$ including r.*

PROOF. Let q be the intersection of L_r and M. Since both q and r are in L_r, there exists a positive number μ which satisfies (7). From positivity of divergences the point q is the unique minimizer of a function $\rho(\cdot, r)$. □

We denote by ρ' the divergence of the dual statistical manifold $(\Omega, \tilde{D}', \tilde{g})$ of $(\Omega, \tilde{D}, \tilde{g})$. Then $\rho(p,q) = \rho'(q,p)$ holds. Therefore, on the same assumption of Theorem 5.1, it follows that

$$\rho'(r,p) = \rho'(r,q) + \mu\rho'(q,p).$$

Recalling that divergences are contrast functions, by virtue of Corollary 5.2, we can call leaves of $\mathscr{F}^{\perp}$ minimum contrast leaves with respect to the dual divergence ρ' [2].

Finally, we give examples of the gradient flow along geodesics relative to the dual connection.

On dynamical systems constrained to flat submanifolds, Fujiwara and Amari showed the following theorem and its applications to engineering.

THEOREM 5.3. ([4, Theorem 2]) *Let $N = \{p_\xi | \xi \in \Xi \subset \mathbf{R}^n\}$ be a submanifold embedded in a flat manifold $\tilde{N}$ with respect to a dualistic structure $(\tilde{\nabla}, \tilde{\nabla}', \tilde{h})$, and (∇, ∇', h) the induced dualistic structure on N. If N is $\tilde{\nabla}$-autoparallel, then for $r \in \tilde{N}$ the gradient flow defined by*

$$\frac{d\xi^i}{dt} = -\sum_{j=1}^{n} h^{ij} \frac{\partial}{\partial \xi^j} \rho(p_\xi, r) \qquad (i = 1, \ldots, n)$$

converges to a unique stationary point independent of the initial point along a ∇'-geodesic, where $\xi = (\xi^1, \ldots, \xi^n)$ is a ∇-affine coordinate such that $\nabla_X(\partial/\partial\xi^j) = 0$ for $X \in \mathscr{X}(N)$, $h_{ij} = h(\partial/\partial\xi^i, \partial/\partial\xi^j)$, $(h^{ij}) = (h_{ij})^{-1}$, and ρ is the divergence of $(\tilde{N}, \tilde{\nabla}, \tilde{h})$. Then the stationary point $q \in M$ is the unique one such that

$$\rho(p,r) = \rho(p,q) + \rho(q,r).$$

A $\tilde{\nabla}$-autoparallel statistical submanifold of a flat statistical manifold $(\tilde{N}, \tilde{\nabla}, \tilde{h})$ is flat, and its divergence coincides with the restriction of the divergence of $(\tilde{N}, \tilde{\nabla}, \tilde{h})$ [1]. These facts imply Theorem 5.3.

We shall investigate a dynamical system constrained to a 1-conformally flat statistical submanifold. Let (M, D, g) be a 1-conformally flat statistical submanifold of an $(n+1)$-dimensional Hessian domain $(\Omega, \tilde{D}, \tilde{g} = \tilde{D}d\varphi)$, where M is an n-dimensional level surface of φ. We assume that (M, D, g) is 1-conformally equivalent to a flat statistical manifold $(M, \bar{D}, \bar{g})$, and that a function ϕ on an open subset U of M satisfies that

$$\bar{g}(X, Y) = e^{\phi} g(X, Y),$$
$$g(\bar{D}_X Y, Z) = g(D_X Y, Z) - d\phi(Z) g(X, Y)$$

for $X, Y, Z \in \mathscr{X}(U)$. We treat an affine coordinate system $\{\bar{x}^1, \ldots, \bar{x}^n\}$ on U such that $\bar{D} d\bar{x}^i = 0$ for $i = 1, \ldots, n$. Let r be a point in a leaf L_q for $q \in U$. We consider $\rho(\cdot, r)$ as a function on U of variables $\bar{x}^1, \ldots, \bar{x}^n$ and denote by $\rho(p_{\bar{x}}, r)$ its value at $p \in U$, where ρ is the divergence of $(\Omega, \tilde{D}, \tilde{g})$. We set $\bar{g}_{ij} = \bar{g}(\partial/\partial\bar{x}^i, \partial/\partial\bar{x}^j)$ for $i,j = 1, \ldots, n$ and $(\bar{g}^{ij}) = (\bar{g}_{ij})^{-1}$ on U, and then we obtain:

COROLLARY 5.4. *The gradient flow defined by*

$$\frac{d\bar{x}^i}{dt} = -\sum_{j=1}^{n} \bar{g}^{ij} \frac{\partial}{\partial \bar{x}^j} \rho(p_{\bar{x}}, r) \qquad (i = 1, \ldots, n) \tag{8}$$

converges to the point q following a D'-geodesic, if U includes the trajectory of the D'-geodesic from an initial point to q.

PROOF. Let μ be the positive number such that $\tilde{\imath}(r) = \mu\tilde{\imath}(q)$. By Theorem 5.1, $\rho(p, r) = \mu\rho(p, q) + \rho(q, r)$ holds. Denoting by $\bar{\rho}$ the divergence of $(M, \bar{D}, \bar{g})$, we have $\bar{\rho}(p, q) = e^{\phi(q)}\rho(p, q)$ by [7]. Thus the gradient flow (8) is equivalent to

$$\frac{d\bar{x}^i}{dt} = -\mu e^{-\phi(q)} \sum_{j=1}^{n} \bar{g}^{ij} \frac{\partial}{\partial \bar{x}^j} \bar{\rho}(p_{\bar{x}}, q) \qquad (i = 1, \ldots, n).$$

Let $\bar{D}'$ be the dual affine connection of $\bar{D}$, and $\{\bar{x}'_1, \ldots, \bar{x}'_n\}$ the dual affine coordinate system of $\{\bar{x}^1, \ldots, \bar{x}^n\}$. Since $\bar{g}^{ij} = \partial\bar{x}'_i/\partial\bar{x}^j$, we have

$$\frac{d\bar{x}'_i}{dt} = -\mu e^{-\phi(q)} \frac{\partial}{\partial \bar{x}^i} \bar{\rho}(p_{\bar{x}}, q) \qquad (i = 1, \ldots, n).$$

Considering the proof of Theorem 5.3 (see [4]), we have

$$\frac{d\bar{x}'_i}{dt} = -\mu e^{-\phi(q)} (\bar{x}'_i - \bar{x}'_i(q)) \qquad (i = 1, \ldots, n).$$

Setting $A = \mu e^{-\phi(q)} > 0$, we obtain

$$\bar{x}'_i = \bar{x}'_i(q) + (\bar{x}'_i(p_0) - \bar{x}'_i(q)) e^{-At} \qquad (i = 1, \ldots, n),$$

where $p_0 \in U$ is an initial point of (8). Thus the flow (8) converges to q following a straight line with respect to a coordinate system $\{\bar{x}'_1, \dots, \bar{x}'_n\}$. Since $\bar{D}'$ is flat and $\bar{D}' d\bar{x}'_i = 0$, the line is a pseudo-geodesic with respect to $\bar{D}'$. From projective equivalence with $\bar{D}'$ and D', a pseudo-geodesic with respect to $\bar{D}'$ is one with respect to D'. Hence the flow (8) converges to q following a D'-geodesic, independent of an initial point. □

References

[1] S. Amari, Differential-Geometrical Methods in Statistics, Lecture Notes in Statistics **28**, Springer-Verlag, 1985.

[2] S. Eguchi, Geometry of minimum contrast, Hiroshima Math. J. **22** (1992), 631–647.

[3] L. P. Eisenhart, Non-Riemannian Geometry, Amer. Math. Soc. Colloq. Publ. **8**, 1927.

[4] A. Fujiwara and S. Amari, Gradient systems in view of information geometry, Physica D **80** (1995), 317–327.

[5] S. Ivanov, On dual-projectively flat affine connections, J. of Geom. **53** (1995), 89–99.

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

[7] T. Kurose, On the divergence of 1-conformally flat statistical manifolds, Tôhoku Math. J. **46** (1994), 427–433.

[8] H. Matsuzoe, On realization of conformally-projectively flat statistical manifolds and the divergences, Hokkaido Math. J. **27** (1998), 409–421.

[9] T. Matumoto, Any statistical manifold has a contrast function—On the C^3-functions taking the minimum at the diagonal of the product manifold, Hiroshima Math. J. **23** (1993), 327–332.

[10] Y. Nakamura, Gradient systems associated with probability distributions, Japan J. Indust. Appl. Math. **11** (1994), 21–30.

[11] K. Nomizu and U. Pinkall, On the geometry and affine immersions, Math. Z. **195** (1987), 165–178.

[12] A. Ohara, N. Suda and S. Amari, Dualistic differential geometry of positive definite matrices and its applications to related problems, Linear Algebra and Its Applications **247** (1996), 31–53.

[13] H. Shima, Geometry of Hessian domains, Differential Geometry and its Applications **7** (1997), 277–290.

[14] H. Shima and J. H. Hao, Geometry associated with normal distributions, Preprint.

[15] K. Uohashi, A. Ohara and T. Fujii, 1-conformally flat statistical submanifolds, to appear in Osaka J. Math.

Keiko Uohashi
Department of Systems and Human Science
Graduate School of Engineering Science
Osaka University
Osaka 560-8531, Japan
keiko@ft-lab.sys.es.osaka-u.ac.jp

Atsumi Ohara
Department of Systems and Human Science
Graduate School of Engineering Science
Osaka University
Osaka 560-8531, Japan
ohara@sys.es.osaka-u.ac.jp

Takao Fujii
Department of Systems and Human Science
Graduate School of Engineering Science
Osaka University
Osaka 560-8531, Japan
fujii@sys.es.osaka-u.ac.jp

Current address of the first-named author: Department of Electrical Engineering and Computer Science, Osaka Prefectural College of Technology, Osaka 572-8572, Japan, uohashi@ecs.osaka-pct.ac.jp