Fundamental Journal of Mathematics and Mathematical Sciences
Vol. 6, Issue 1, 2016, Pages 51-66
This paper is available online at http://www.frdint.com/
Published online September 4, 2016

ON GRADIENT AND HAMILTONIAN FLOWS ON EVEN DIMENSIONAL DUALLY FLAT SPACES

NOBUTAKA BOUMUKI and TOMONORI NODA*

Department of Mathematics
Tokyo University of Science
1-3 Kagurazaka, Shinjuku-ku
Tokyo 162-8601, Japan
e-mail: boumuki@rs.tus.ac.jp

General Education and Research Center
Meiji Pharmaceutical University
2-522-1 Noshio, Kiyose-shi
Tokyo 204-8588, Japan
e-mail: noda@my-pharm.ac.jp

Abstract

In this paper, we completely elucidate Fujiwara's results which assert that some kind of gradient flow on any even dimensional dually flat space can be expressed as Hamiltonian flow by using Darboux type theorem for the canonical symplectic structure on the cotangent bundle of the space. This gives a variational characterization of the Fisher metric.

Keywords and phrases: special Kähler, dually flat space, Darboux's theorem, gradient flow, Hamiltonian flow.

2010 Mathematics Subject Classification: Primary 53A15; Secondary 53D25.

*Corresponding author

Received July 25, 2016; Accepted August 8, 2016

1. Introduction

Let (S, g, ∇) be a statistical manifold. Namely, g is a (pseudo) Riemannian metric and ∇ is a torsion free affine connection on differentiable manifold S such that ∇g is symmetric. Statistical manifolds are abstract of statistical models and statistical inferences, for instance, on exponential families, e- and m- connections correspond to maximum likelihood estimations. In particular, flat cases are very important and are called *dually flat spaces* (or *Hessian manifolds*). On a dually flat space, there exist dual structures as dual connections ∇ and ∇^*, dual coordinates $\{\theta\}$ and $\{\eta\}$, and potential functions $\varphi(\theta)$ and $\psi(\eta)$. Moreover, by using potential functions, we can construct divergence functions (or contrast functions) on the square of the space. The relative entropy (or a KL divergence) on an exponential family is an example of divergence functions.

In Fujiwara [8, 9], the following properties of gradient equations with a divergence as a potential function are studied. On a dually flat space (S, g, ∇, ∇^*), fix a point q in S. Define $U(p) := D_\nabla(p, q)$ and consider the gradient flow equation $\dot{\theta} = -\mathrm{grad}\, U(\theta)$, so that in η coordinate we have

$$\dot{\eta} = -(\eta - \eta(q)), \tag{1.1}$$

and the solution is given by $\eta(p(t)) = \eta(q) + (\eta(p(0)) - \eta(q))e^{-t}$. This solution converges to q along an m-geodesic. Replacing D_∇ in U by D_{∇^*}, we obtain a steepest ascent flow of entropy. These results also hold for constraint systems on exponential families. Moreover, it is shown that gradient flow equations as in (1.1) can be expressed as Hamiltonian equations when the space has even dimension. Namely, if $\dim S = 2n$, then a gradient flow equation $\dot{\eta} = -\eta$ on S coincides with a Hamiltonian equation with position $Q_k = \eta_{2k}$, momentum $P^k = -1/\eta_{2k-1}$ and Hamiltonian $H = -Q_k P^K$. This result which shows a relationship between Hamiltonian and gradient flows is very mysterious and interesting because aspects of Hamiltonian flow are different from that of gradient flow in general. The purpose of this article is to make clear theoretical reasons of this relationship from symplectic geometric view point.

A symplectic structure is a differential 2 form satisfying nondegeneracy and integrability conditions on a differentiable manifold, and symplectic manifold is an abstract of phase spaces in classical mechanics. To unravel phenomena due to Fujiwara, we have to clear the symplectic structure on appropriate space, and in this paper, we give 3 kinds of correspondences between gradient and Hamiltonian flows on even dimensional dually flat spaces. In first and second ones, we use a Darboux type theorem for symplectic structure (indeed, special Kähler structure) on the square of a special dually flat space. To use symplectic reduction argument is the third.

On exponential family *S*, natural coordinates $\{\theta^i\}$ and expectation coordinates $\{\eta_j\}$ are both coordinate systems on *S*, but in fact, these are different world living in. On dually flat spaces not necessarily exponential families, these coordinates can be considered coordinates on the space by the flatness condition. However, we should consider that natural coordinates $\{\theta^i\}$ are on the tangent spaces and expectation coordinates $\{\eta_j\}$ are on the cotangent spaces of the space, and the Fisher metric gives the connection between these coordinates as in the Riesz representation theorem. From this viewpoint, we can associate the canonical symplectic structure to the square of the dually flat space, and then we can understand the phenomena for gradient flows above.

To see the relation between gradient and Hamiltonian flows on dually flat space, it is important that the existence of special dually flat space, say *L*. Considering the canonical symplectic structure on $L \times L$, we see the reason why gradient flow equations (1.1) correspond to Hamiltonian equations.

Theorem 1.1. *Let S be an even dimensional dually flat space. For any point on S, there exist a local neighborhood U of the point and a dually flat space L such that U is isomorphic to* $L \times L$ *as statistical manifold. The Hamiltonian equation whose Hamiltonian is canonical divergence with respect to the canonical symplectic structure on* $L \times L$ *coincides with gradient flow equation* (1.1).

The key of this theorem is that both gradient flow (1.1) and Hamiltonian flow on even dimensional dually flat space can be expressed as *m*-geodesic, where is a generalization of a result “a geodesic is critical point of both length and energy functions” in Riemannian geometry. Hence we obtain a variational characterization of Fisher metrics on even dimensional dually flat spaces.

It should be able to see a part of connections among information geometry, classical and quantum physics. Indeed, Goto [12, 13] and Mrugała [19] study connections between thermodynamics and contact geometry which is odd dimensional analogue of symplectic geometry; Molitor [18] describes a connection between information geometry and quantum mechanics; connections between symplectic geometry and quantum mechanics are given by de Gosson [10], de Gosson and Luef [11], Censi [5] and Hsiao and Scheeres [15].

2. Statistical Manifolds and Gradient Flow Equations

In this section, we review some results of information geometry and Fujiwara [8]. For general theory of information geometry, see Amari and Nagaoka [3], Eguchi [6], Matsuzoe [17], Shima [23] and the references therein.

For dually flat space (S, g, ∇, ∇^*), there exist ∇-affine coordinates $(\theta^1, \dots, \theta^n)$ and ∇^*-affine coordinates $(\eta_1, \dots, \eta_n)$ so that

$$g\left(\frac{\partial}{\partial\theta^i}, \frac{\partial}{\partial\eta_j}\right) = \delta^i_j. \tag{2.1}$$

We refer $(\theta^1, \dots, \theta^n)$ and $(\eta_1, \dots, \eta_n)$ to dual coordinates and set $\partial_i := \partial/\partial\theta^i$, $\partial^j := \partial/\partial_{\eta_j}$. Moreover, there exist (local) functions $\psi, \varphi : S \to \mathbb{R}$ satisfying $d\psi = \eta_i d\theta^i$, $d\psi = \theta^i d\eta_i$ such that $\varphi + \psi = \theta^i\eta_i$. We collect relations between objects defined above as a lemma.

Lemma 2.1. *Let* $(\theta^1, \dots, \theta^n)$ *and* $(\eta_1, \dots, \eta_n)$ *be dual coordinates satisfying* (2.1) *on a dually flat space. Then the following hold*

$$\theta^i = \frac{\partial\varphi}{\partial\eta_i}, \eta_j = \frac{\partial\psi}{\partial\theta^j}, d\theta^i = \frac{\partial\theta^i}{\partial\eta_j} d\eta_j = g^{ij} d\eta_j, d\eta_j = \frac{\partial\eta_j}{\partial\theta^i} d\theta^i = g_{ij} d\theta^i,$$

$$g_{ij} := g(\partial_i, \partial_j) = \frac{\partial\eta_j}{\partial\theta^i} = \frac{\partial^2\psi}{\partial\theta^i\partial\theta^j}, g^{ij} := g(\partial^i, \partial^j) = \frac{\partial\theta^i}{\partial\eta_j} = \frac{\partial^2\varphi}{\partial\eta_i\partial\eta_j}.$$

We now define the ∇^*-divergence $D_{\nabla^*} : S \times S \to \mathbb{R}$ by

$$D_{\nabla^*}(p, q) := D(p\|q) := \varphi(p) + \psi(q) - \eta_i(p)\theta^i(q).$$

We also define ∇-divergence by interchanging p and q in D_{∇^*}, that is, $D_{\nabla}(p\|q) := D_{\nabla^*}(q\|p)$. Then D_{∇} and D_{∇^*} are independent of the choice of affine coordinates and define functions on $S \times S$. For any (p, q) in $S \times S$, $D_{\nabla}(p\|q) \geq 0$ and the equality holds if and only if $p = q$, and so does D_{∇^*}. In the case where exponential families $S = \{p_\theta = \exp(C(x) + \theta^i F_i(x) - \psi(\theta))\}$, the divergence

$$D_{\nabla}(p\|q) = \int p \log \frac{p}{q}\, dx$$

is the relative entropy (or Kullback-Leibler divergence). These two divergences are contrast functions defined as follows.

Let S be a differentiable manifold and let $D : S \times S \to \mathbb{R}$ be a non-negative function satisfying identity of indiscemibles, namely $D(p\|q) \geq 0$ for any $(p, q) \in S \times S$ and the equality holds if and only if $p = q$. If we define a matrix $g^D = [g_{ij}^D]$ by

$$g_{ij}^D := D[\partial_i \partial_j \|\cdot] = D[\cdot\|\partial_i \partial_j] = -D[\partial_i \|\partial_j],$$

then $g^D = [g_{ij}^D]$ is positive semi-definite, where $D[\cdot\|\cdot]$ denotes the restriction of D to the diagonal of $S \times S$. A function D on $S \times S$ is a contrast function if matrix g^D is positive-definite, and then g^D is a Riemannian metric on S. As we saw above, a statistical structure naturally defines contrast functions, conversely a contrast function defines statistical structures on the manifold.

We next refer to results of gradient flow equations on dually flat spaces due to Fujiwara [8, 9]. Note that the notations of divergence functions in [8] are reverse to that of our divergences.

Theorem 2.2. *Let (S, g, ∇, ∇^*) be a dually flat space, and fix a point q in S.*

(1) *Define a function on S by* $U(p) := D_{\nabla}(p, q)$, *and consider the gradient flow equation*

$$\dot{\theta} = -\mathrm{grad}\, U(\theta). \tag{2.2}$$

Then the solution of this equation is given by

$$\eta(p(t)) = \eta(q) + (\eta(p(0)) - \eta(q))e^{-t}$$

which converges to q along $m*$*-geodesic as* $t \to \infty$.

(2) *If S has even dimension*, *then gradient flow equation*

$$\dot{\eta} = -\eta \tag{2.3}$$

can be reformulated as a Hamiltonian equation whose position is $Q_k = \eta_{2k}$, *momentum is* $P^k = -1/\eta_{2k-1}$, *and Hamiltonian is* $H = -Q_k P^k$ $(k = 1, \ldots, n)$.

Proof. (1) Since $D(p\|q) = \psi(p) - \psi(q) + (\theta(q) - \theta(p)) \cdot \eta(q)$, we have $\partial_j U(\theta) = \partial_j \psi(p) - \eta_j(q)$, and

$$\dot{\theta}^i = -g^{ij}(\eta_j - \eta_j(q))$$

because $\partial_j \psi = \eta_j$. By multiplying g_{ji} the both sides of the equality and taking the summation, we have $g_{ji}\dot{\theta}^i = -(\eta_j - \eta_j(q))$. On the other hand

$$g_{ji}\dot{\theta}^i = \frac{\partial \eta_j}{\partial \theta^i}\frac{d\theta^i}{dt} = \frac{d\eta_j}{dt},$$

so that in η-coordinates the gradient flow equation (2.2) has form

$$\dot{\eta} = -(\eta - \eta(q)).$$

By integrating this equality, we obtain $\eta(p(t)) = \eta(q) + (\eta(p(0)) - \eta(q))e^{-t}$.

(2) In view of (2.3), we have

$$\frac{d}{dt}H = \frac{1}{\eta_{2k-1}^2}(\dot{\eta}_{2k}\eta_{2k-1} - \eta_{2k}\dot{\eta}_{2k-1}) = 0.$$

Direct calculation gives us the equivalence of the Hamiltonian equation

$$\frac{dQ_k}{dt} = \frac{\partial H}{\partial P^k}, \quad \frac{dP^k}{dt} = -\frac{\partial H}{\partial Q_k}$$

and

$$\dot{\eta}_{2k} = -\eta_{2k}, \quad \dot{\eta}_{2k-1} = -\eta_{2k-1}.$$

For $f \in \mathcal{F}$, vector field grad f is defined by $g(\text{grad } f, Y) = df(Y)$, and is expressed as $(g^{ij}\partial_j f)$ in coordinates. Hence gradient flow equation (2.2) is expressed as

$$\dot{\theta}^i = -g^{ij}\partial_j U(\theta), \quad \dot{\eta} = -(\eta - \eta(q))$$

in θ (∇ affine), η (∇^* affine) coordinates, respectively. In the case where normal distributions, if we take delta distribution at the origin as q, then we get a result in Nakamura [20]. In the case where exponential families, the potential function for ∇ divergence is $U(\eta) = -(\text{entropy of } p) + \text{const.}$, then the gradient flow equation is a steepest ascent flow of entropy. Especially by taking a uniform distribution as q it becomes an Ornstein and Uhlenbeck process [22].

While (2) of the theorem shows that gradient flow equation (2.3), more generally (2.2) is reformulated as Hamiltonian equation, it is very mysterious and interesting result. In the next section, we clear the reason why gradient flow can be rewritten to Hamiltonian equation.

3. Symplectic Structure and Hamiltonian Equations on T^*S

First of all, we construct symplectic structures on the square of dually flat space from contrast functions. To distinguish between the first and second factors of $S \times S$, we denote not $S \times S$ but $S_1 \times S_2$ and put * on coordinates of the second factor, e.g., ξ_i^*, θ_j^* and so on. Let $D : S_1 \times S_2 \to \mathbb{R}$ be a contrast function, and define a map form $S_1 \times S_2$ to T^*S by

$$d_1 D : S_1 \times S_2 \to T^* S$$

$$(\xi, \xi^*) \mapsto (\xi, d_1 D),$$

where $d_1 D = \frac{\partial D}{\partial \xi^i} d\xi^i$ denotes the exterior derivative along the first factor of $S_1 \times S_2$. Define a two form ω by the pull-back of the canonical symplectic structure on $T^* S$ with respect to this map $d_1 D$:

$$\omega(\xi, \xi^*) := (d_1 D)^* \ (-d\theta_0) = \frac{\partial^2 D}{\partial \xi^i \partial \xi^{*j}} d\xi^i \wedge d\xi^{*j}.$$

This ω is a symplectic form on a neighborhood of the diagonal of $S_1 \times S_1$. In general, ω is only nondegenerate on a neighborhood of the diagonal of $S_1 \times S_2$, and then not necessarily define a global symplectic structure on $S_1 \times S_2$. However, we do not need the range where ω is nondegenerate in this paper, from now on we assume that ω above define a global symplectic form on $S_1 \times S_2$. For detail see Barndorff-Nielsen and Jupp [4] and Noda [21].

For example, if we take $S = \mathbb{R}^n$ and $D : S_1 \times S_2 \to \mathbb{R}$ as

$$D(\xi \| \xi^*) = \frac{1}{2} \|\xi - \xi^*\|^2,$$

where $\|\cdot\|$ denotes the Euclid norm. Then this is a contrast function on $\mathbb{R}^n \times \mathbb{R}^n$, and the induced symplectic structure is the canonical symplectic structure

$$\omega_0 = \sum_i d\xi^i \wedge d\xi^{*i} = \delta_{ij} d\xi^i \wedge d\xi^{*j}$$

on $T^* \mathbb{R}^n = \mathbb{R}^{2n}$. Hence the symplectic structure associate with the Hellinger distance is the canonical symplectic structure on the cotangent bundle.

For general dually flat spaces, we have

Lemma 3.1. *Let* (S, g, ∇, ∇^*) *be a dually flat space. Then the symplectic structure on* $S_1 \times S_2$ *defined by* D_∇ *is given by*

$$\omega_0 = d\eta_i^* \wedge d\theta^i = -g_{ij}^* d\theta^i \wedge d\theta^{*j} = -g_{ij}^* g^{ik} d\eta_k \wedge d\theta^{*j} = -g^{ij} d\eta_i \wedge d\eta_j^*.$$

If we use D_{∇^*}, *then*

$$\omega_0^* = -d\eta_i \wedge d\theta^{*i} = -g_{ij} d\theta^i \wedge d\theta^{*j} = -g_{ik} g^{*kj} d\theta^i \wedge d\eta_j^* = -g^{*ij} d\eta_i \wedge d\eta_j^*.$$

Proof. Let $(\theta^1, \dots, \theta^n)$ and $(\eta_1, \dots, \eta_n)$ be ∇ affine and ∇^* affine coordinates on S, respectively. By choosing (θ^i, η_j^*) as coordinates on $S_1 \times S_2$, we have

$$\frac{\partial^2 D_\nabla}{\partial \theta^i \partial \eta_j^*} = -\delta_i^j,$$

and in view of Lemma 2.1, we obtain the results.

Note that in expressions of ω_0 above, the second term $\omega_0 = -g_{ij}^* d\theta^i \wedge d\theta^{*j}$ is symplectic structure, more precisely Kähler structure, defined on the tangent bundle of S. Moreover, we can define a special Kähler structure on $S_1 \times S_2$ by setting $g_\theta \oplus g_{\eta^*}$ and $\nabla \oplus \nabla^*$. Then the associate complex structure J satisfies $J : \partial_i = -\partial^{*i}$. (For special Kähler manifolds, see Freed [7], Alekseevsky-Cortés-Devchand [2].) While we can define symplectic structures on $S_1 \times S_2$ by $d\theta^i \wedge d\theta^{*i}$ and $d\eta_i \wedge d\eta_i^*$, we ignore these structures since they does not induced from divergence functions.

Now we see properties of symplectic structure ω_0 as in Lemma 3.1. First of all, we see alternative way to get ω_0 which gives a classical mechanical characterization of the structure.

Consider coordinates $(\eta_1^*, \dots, \eta_n^*)$ on $S_1 \times S_2$. Namely, if we regard the first factor S_1 as tangent space and second factor S_2 as cotangent space, and $S_1 \times S_2$ as the cotangent bundle of S, then the canonical symplectic structure on $S_1 \times S_2$ is given by

$$\omega_0 = d\eta_i^* \wedge d\theta^i,$$

where $(\theta^1, \dots, \theta^n)$ and $(\eta_1^*, \dots, \eta_n^*)$ are ∇ affine and ∇^* affine coordinates on

S_1 and S_2, respectively. This symplectic structure coincides with the form as in Lemma 3.1, so we use the same notation. From this viewpoint, potential functions ψ and φ^* satisfy a Legendre relation in analytical mechanics.

Lemma 3.2. *On symplectic manifold* $(S_1 \times S_2, \omega_0)$, *the Hamiltonian vector field of the canonical divergence* D_∇ *is given by*

$$X_{D_\nabla}(p, q) = (\theta^i(p) - \theta^{*i}(q))\frac{\partial}{\partial \theta^i} - (\eta_i^*(q) - \eta_i(p))\frac{\partial}{\partial \eta_i^*}. \tag{3.1}$$

Proof. For the canonical divergence $D_\nabla(p\|q) = \varphi^*(q) + \psi(p) - \eta_i^*(q)\theta^i(p)$, we have

$$\begin{aligned} dD_\nabla &= \frac{\partial \varphi^*}{\partial \eta_i^*} d\eta_i^* + \frac{\partial \psi}{\partial \theta^i} d\theta^i - \theta^i d\eta_i^* - \eta_i^* d\theta^i \\ &= \theta^{*i} d\eta_i^* + \eta_i d\theta^i - \theta^i d\eta_i^* - \eta_i^* d\theta^i \\ &= (\theta^{*i} - \theta^i) d\eta_i^* + (\eta_i - \eta_i^*) d\theta^i. \end{aligned}$$

It follows that the first factor of X_{D_∇} generates ∇-geodesic, and the second generates ∇^*-geodesic. If we fix a point in S_2 and identify S_2 with S by using a certain parallel translation, then the Hamiltonian vector field X_{D_∇} reduces to

$$(\theta^i(p) - \theta^i(q))\frac{\partial}{\partial \theta^i},$$

hence this vector filed on S generates ∇-geodesic. Similarly, fixing a point in S_1 and taking a parallel translation, we get a vector field

$$(\eta_i(q) - \eta_i(p))\frac{\partial}{\partial \eta_i}$$

on S which generates ∇^*-geodesic. Therefore geodesics on S with respect to dual connections can be obtained by Hamiltonian flow for the divergence.

Under these preparations, we describe the reason why gradient flow equation

(1.1) can be expressed as a Hamiltonian equation. While we notice that (1.1) looks like the second factor of (3.1), at this time the symplectic structure is unclear. In this reason, we construct a dually flat space L such that $(L \times L, \omega_0)$ is (locally) symplectic diffeomorphic to S and a Hamiltonian flow on $L \times L$ coincides with gradient flow (1.1).

In Lemma 3.2, we calculate Hamiltonian vector filed X_{D_∇} on $S \times S$ with respect to ω_0, in view of the definition of the divergence D_∇, for example, we try to compute the Hamiltonian vector field of $H = -\eta^*(p)\theta(q)$ which is a part of the divergence. Then we have

$$X_H = \theta^i \frac{\partial}{\partial \theta^i} - \eta_i^* \frac{\partial}{\partial \eta_i^*},$$

and we want to show this coincides with (2.3).

Lemma 3.3. *Let* L^m *be a dually flat space defined as the following*: $L = \{\theta^1, \ldots, \theta^m\}$, $\psi(\theta) = -\sum_i (1 - \log|\theta^i|)$, $\eta^i = -(\theta^i)^{-1}$ *and* $\varphi(\eta) = -\sum \log|\eta_i|$. *For this* L, *the symplectic structure associated to divergence* D_∇ *on* $(L_1 \times L_2, \omega_0)$ *is given by*

$$\omega_0 = d\eta_i^* \wedge d\theta^i = \sum_i \frac{d\eta_i^* \wedge d\eta_i}{\eta_i^2},$$

and the Hamiltonian vector filed of $H = -\eta_i^* \theta^i$ *is*

$$X_H = -\eta_i \frac{\partial}{\partial \eta_i} - \eta_i^* \frac{\partial}{\partial \eta_i^*}.$$

Proof. By $\eta_i = -(\theta^i)^{-1}$, we have $g^{ij} = \eta_i^2 \delta^{ij}$, and then

$$d\eta_i^* \wedge d\theta^i = d\eta_i^* \wedge g^{ij} d\eta_j = \sum_i \frac{d\eta_i^* \wedge d\eta_i}{\eta_i^2}.$$

Next for $H = -\eta_i^* \theta^i$, we have

$$X_H = \theta^i \frac{\partial}{\partial \theta^i} - \eta_i^* \frac{\partial}{\partial \eta_i^*},$$

by using $\theta^i = -\eta_i^{-1}$ and

$$\frac{\partial}{\partial \theta^i} = \frac{\partial \eta_j}{\partial \theta^i} \frac{\partial}{\partial \eta_j} = \eta_i^2 \frac{\partial}{\partial \eta_i},$$

we obtain

$$X_H = -\eta^i \frac{\partial}{\partial \eta^i} - \eta_i^* \frac{\partial}{\partial \eta_i^*}.$$

It follows Darboux type theorem for even dimensional dually flat spaces.

Theorem 3.4. *Let S be a $2n$ dimensional dually flat space. For any point $p \in S$, there exist a neighborhood U of p and canonical symplectic structure*

$$\omega_1 = \sum_{k=1}^{n} \frac{d\eta_{2k-1} \wedge d\eta_{2k}}{\eta_{2k-1}^2}$$

on U such that

$$(L^n \times L^n, \omega_0) \simeq (U, \omega_1), \tag{3.2}$$

where $\simeq$ means symplectic diffeomorhic.

This theorem shows that on any dually flat space with even dimension there is a natural symplectic structure, and gradient flow equation (2.3) coincides with Hamiltonian equation with $H = -\eta^* \theta$ because of the lemma above. Hence, we obtain a correspondence between gradient and Hamiltonian flow on dually flat spaces. By Theorem 3.4, if S has even dimension, then any dually flat space admits a natural symplectic structure defined from the canonical divergence which associate to Poincaré metric. Hence, we should consider Theorem 3.4 as Darboux theorem on dually flat spaces (indeed, this is for special Kähler structures).

This is the first method, but is hard to be able to see a connection between the structures. So we give a supplement of the relation between symplectic structure ω_1

and statistical structure on the space. Fix a point q in $2n$ dimensional dually flat space S, and let $U(p) = D_{\nabla}(p\|q)$. Then for $1 \le i \le 2n$, we have

$$\partial_i U = \eta_i - \eta_i(q).$$

We separate $\eta_1, \dots, \eta_{2n}$ into even and odd indexes, and take even indexes as η coordinates and odd indexes as $-\theta^{-1}$ coordinates which think of that of cotangent spaces. In this case, expressing the symplectic structure ω_0 by η coordinates, then we obtain ω_1. Moreover, in a direction gradient flow is expressed as *m*-geodesic flow, in another direction Hamiltonian flow is *m*-geodesic flow.

Newt we see more direct correspondence between two kinds of geometric flows. As we see above, the second factor of the Hamiltonian vector field $X_{D_{\nabla}}$ looks like the gradient flow $\dot{\eta} = (\eta_i(q) - \eta_i(p))$ which is expressed by η-coordinates on S. To identify S with the second factor of $S_1 \times S_2$ by using a parallel translation, we take L as in Lemma 3.2 instead of S, where we need even dimensional L. We consider L^m such that $m = 2n = \dim S$, then the second part of Hamiltonian vector field $X_{D_{\nabla}}$ generates an *m*-geodesic by Lemma 3.2, and then corresponds to (2.3).

In these correspondences, the essential reason is that potentials of Fisher metric on dually flat space S and of symplectic structure on $S \times S$ are both canonical divergences, and geodesics can be cahracterized to Riemannian and symplectic geometry.

As the last, we construct a correspondence between gradient and Hamiltonian equations via the symplectic reduction argument. For detail of symplectic reductions, see Abraham and Marsden [1], Guillemin and Sternberg [14] and Marsden and Weinstein [16].

Although, we also consider the space as in Lemma 3.2, we need even dimensional space, say $2n$, so we use the notation S' instead of L. Namely, let S' satisfies $S' = \{\theta^1, \dots, \theta^{2n}\}$, $\psi(\theta) = -\sum_i (1 - \log|\theta^i|)$, $\eta^i = -(\theta^i)^{-1}$ and $\varphi(\eta) = -\sum \log|\eta_i|$. We choose $(\theta^1, \dots, \theta^{2n}, \eta_1^*, \dots, \eta_{2n}^*)$ as a coordinates system on $S_1' \times S_2'$. Define the action of $\mathbb{R}^n$ on $S_1' \times S_2'$ by

$$a \cdot (\theta, \eta^*) := (\ldots, \theta^{2k-1}, \theta^{2k}, \ldots, ; \ldots \eta^*_{2k-1} + a_k, \eta^*_{2k} + a_k, \ldots)$$

for every $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$. The infinitesimal generator is $V_k = \partial^{*2k-1} + \partial^{*2k}$, $1 \leq k \leq n$, and because

$$i(V_k)\omega_0 = d(\theta^{2k-1} + \theta^{2k}) = 0,$$

the moment map of this action is $(\theta^1 + \theta^2, \ldots, \theta^{2n-1} + \theta^{2n})$. Moreover, since the reduced space at the origin can be identified with the subset $\eta^*_{2k-1} + \eta^*_{2k} = 0$ in $\mu^{-1}(0)$, the reduced symplectic structure $\tilde{\omega}_0$ coincides with ω_0 restricted to

$$\{(\theta, \eta^*); \theta^{2k-1} + \theta^{2k} = 0, \eta^*_{2k-1} + \eta^*_{2k} = 0, 1 \leq k \leq n\}.$$

Hence, we have

$$\tilde{\omega}_0 = 2\sum_{k=1}^{n} d\eta^*_{2k-1} \wedge d\theta^{2k-1} = 2\sum_{k=1}^{n} d\eta^*_{2k-1} \wedge g^{2k-1, j} d\eta_j$$

$$= 2\sum_{k=1}^{n} \frac{d\eta^*_{2k-1} \wedge d\eta_{2k-1}}{\eta^2_{2k-1}}$$

which coincides with the symplectic structure in Theorem 3.4 up to scalar multiplication. Note that under the identification of the reduced space, we have

$$\frac{\theta^1}{\eta^*_1} = \cdots = \frac{\theta^{2n}}{\eta^*_{2n}}.$$

References

[1] R. Abraham and J. E. Marsden, Foundations of Mechanics, 2nd ed., Addison-Wesley Publishing Company, Inc., 1987.

[2] D. V. Alekseevsky, V. Cortés and C. Devchand, Special complex manifolds, J. Geom. Phys. 42 (2002), 85-105.

[3] S. Amari and H. Nagaoka, Methods of information geometry, Translations of Mathematical Monographs, 191, Translated from the 1993 Japanese original by Daishi

Harada, American Mathematical Society, Providence, RI; Oxford University Press, Oxford, 2000.

[4] O. E. Barndorff-Nielsen and P. E. Jupp, Statistics, yokes and symplectic geometry, Annales de la faculté des sciences de Toulouse Sér. 6(6) (1997), 389-427.

[5] A. Censi, Symplectic camels and uncertainty analysis, preprint.

[6] S. Eguchi, Geometry of minimum contrast, Hiroshima Math. J. 22 (1992), 631-647.

[7] D. Freed, Special Kähler manifolds, Comm. Math. Phys. 203 (1999), 31-52.

[8] A. Fujiwara, Dynamical systems on statistical models (State of art and perspectives of studies on nonlinear integrable systems), RIMS Kokyuroku 822 (1993), 32-42.

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

[10] M. de Gosson, The symplectic camel and the uncertainty principle, The tip of an iceberg?, Foundations of Physics 99 (2009), 194-214.

[11] M. de Gosson and F. Luef, Symplectic capacities and the geometry of uncertainty: The irruption of symplectic topology in classical and quantum mechanics, Physics Reports 484 (2009), 131-179.

[12] S. Goto, Legendre submanifolds in contact manifolds as attractors and geometric nonequilibrium thermodynamics, J. Math. Phys. 56 (2015), 073301.

[13] S. Goto, Contact geometric descriptions of vector fields on dually flat spaces and their applications in electric circuit models and nonequilibrium statistical mechanics, arXiv:1512.00950v1.

[14] V. Guillemin and S. Sternberg, Symplectic Techniques in Physics, Cambridge University Press, 1986.

[15] F.-Y. Hsiao and D. J. Scheeres, Fundamental constraints on uncertainty relations in Hamiltonian systems, IEEE Transactions in Automatic Control 52 (2007), 686-691.

[16] J. Marsden and A. Weinstein, Reduction of symplectic manifolds with symmetry, Rep. Math. Phys. 5 (1974), 121-130.

[17] H. Matsuzoe, Differential Geometrical Foundations of Information Geometry: Geometry of Statistical Manifolds and Divergences, World Scientific Pub. Co., Inc., 2016.

[18] M. Molitor, Exponential families, Kähler geometry and quantum mechanics, J. Geom. Phys. 70 (2013), 54-80.

[19] R. Mrugała, On contact and metric structures on thermodynamic spaces, RIMS Kokyuroku 1142 (2000), 167-181.

[20] Y. Nakamura, Completely integrable gradient systems on the manifolds of Gaussian and multinomial distributions, Japan J. Indust. Appl. Math. 10 (1993), 179-189.

[21] T. Noda, Symplectic structures on statistical manifolds, J. Aust. Math. Soc. 90 (2011), 371-384.

[22] T. Obata, H. Hara and K. Endo, Differential geometry of nonequilibrium processes, Phys. Rev. A 45 (1992), 6997-7001.

[23] H. Shima, The Geometry of Hessian Structures, World Scientific Publishing, Singapore, 2007.