Information geometry in a global setting

Atsuhide MORI

Abstract. We begin a global study of information geometry. In this article, we describe the geometry of normal distributions by means of positive and negative contact structures associated to the suspension Anosov flows on Sol^3-manifolds.

1. Introduction

Information geometry is basically a local study of a parameter space of a family of probability distributions by means of differential geometry (see [1]). We begin a global study of this object. In this article, we restrict ourselves to the case of normal distributions. Then the parameter space is the upper half-plane $\mathbb{H} = \{(m, s) \in \mathbb{R}^2 \mid s > 0\}$ with a certain non-Kähler metric and a non-metric connection, where m denotes the mean and s the standard deviation. The background of the metric and the connection is a certain potential function D on $\mathbb{H} \times \mathbb{H}$. On the other hand, we know that the product $\mathbb{H} \times \mathbb{H}$ is for parametrizing abelian surfaces. Precisely, the quotient of the product $\mathbb{H} \times \mathbb{H}$ under the action of a Hilbert modular group is a non-compact singular complex surface which parametrizes the isomorphism classes of complex abelian surfaces with a real multiplication structure. Its topology was studied by Hirzebruch [4]. He took the quotient M of $\mathbb{H}\times\mathbb{H}$ under the action of a semi-direct product $\mathbb{Z}\ltimes\mathbb{Z}^2$ of lattices of $\mathbb{R}$ and $\mathbb{R}^2$ which acts on $\mathbb{H}\times\mathbb{H}$ freely and properly discontinuously. Then the neighborhoods of ∞ form the end model for a Hilbert modular surface. From contact topological point of view, $M \cup \{\infty\}$ is the cone with positive and negative symplectic structures whose base M_0 is a Sol^3-manifold with the positive and negative contact structures associated to the suspension Anosov flow (see [6] and §4 below for the precise meaning). Now the result of this article can be summarized as follows.

i) (The model in §3.) We describe the information geometry of $\mathbb{H}$ by introducing a self-correspondence $F \subset \mathbb{H} \times \mathbb{H}$ which identifies the convolution operation for the probability densities in each factor of $\mathbb{H}\times\mathbb{H}$ to the Bayesian learning for those in the other factor.

ii) (The propositions in §5.) The action of $\mathbb{Z} \ltimes \mathbb{Z}^2$ preserves the metric and the connection on each factor of $\mathbb{H} \times \mathbb{H}$. The sum (resp. the difference) of the natural area forms descends to the positive (resp. negative) symplectic structure on the quotient M. The surface F is

2010 *Mathematics Subject Classification.* 57R17, 57R30, 62B10.
Key words and phrases. Information geometry, contact structure, foliation.

a Lagrangian correspondence which descends to a densely immersed Lagrangian submanifold $L(\subset M_0) \subset M$ with respect to the negative symplectic structure. Further L is decomposed into Legendrian submanifolds of M_0 with respect to the positive contact structure.

iii) (Theorem 1.) We can take a contact Hamiltonian flow on the preimage of M_0 such that
   a) the surface F is an invariant submanifold,
   b) the push-forward of the induced flow on F to each of the factors of $\mathbb{H} \times \mathbb{H}$ is tangent to a foliation whose leaves are geodesics for the non-metric connection, and
   c) the iterations of convolutions along the foliation of the first factor correspond to those of Bayesian learnings along the foliation of the second factor under F.

Note that we could take a similar contact flow just partially on the quotient. Indeed the contact Hamiltonian function for the flow in c) is the inverse coefficient of variance m/s on the first factor which is not preserved under the $\mathbb{Z}^2$-action (but preserved under the $\mathbb{Z}$-action). We raise open problems concerning the geometric characterization of the potential function D on $\mathbb{H} \times \mathbb{H}$ and the relation between abelian varieties and pairs of normal distributions (§6).

2. Information geometry

In a smooth setting of parametric statistics, we consider an open set $U \subset \mathbb{R}^n$ and a positive function p on $\mathbb{R} \times U$ with the normality condition

$$\int_{-\infty}^{\infty} p(x, X)dx \equiv 1.$$

Each point $X \in U$ presents a random variable with probability density $p_X(x) = p(x, X)$. We consider the relative entropy

$$D(X, Y) = \int_{-\infty}^{\infty} p_X(x) \left(\log \frac{1}{p_Y(x)} - \log \frac{1}{p_X(x)} \right) dx$$

for $(X, Y) \in U \times U$ as a fundamental potential. It defines a separating premetric [1] on U, which is called the Kullback-Leibler divergence in information geometry. Using the tensor notation only in this section, we define the Fisher metric g and

[1] A premetric or prametric on a set U is a non-negative function on $U \times U$ which vanishes along the diagonal set. If it is positive elsewhere, we say that it is separating. Each value of a premetric is called a distance.

the e(xponential)-connection ∇ by the cubic approximation

$$\begin{aligned} D(X, X+\Delta X) &\approx \frac{1}{2!}\sum_{i,j} g_{ij}\Delta X^i \Delta X^j \\ &\quad + \frac{1}{3!}\sum_{i,j,k}(\partial_i g_{jk} + \Gamma_{jk,i})\Delta X^i \Delta X^j \Delta X^k \end{aligned}$$

$$\Bigl(\text{or}\quad \begin{aligned} D(X+\Delta X, X) &\approx \frac{1}{2!}\sum_{i,j} g_{ij}\Delta X^i \Delta X^j \\ &\quad + \frac{1}{3!}\sum_{i,j,k}(\partial_i g_{jk} + \partial_j g_{ki} - \Gamma_{ij,k})\Delta X^i \Delta X^j \Delta X^k\Bigr) \end{aligned}$$

of a small distance, where the coefficients in the expression(s) are supposed to be symmetric, eg., $\partial_i g_{jk} + \Gamma_{jk,i} = \partial_j g_{ki} + \Gamma_{ki,j}$. The coefficients $\Gamma_{ij,k}$ of the e-connection all vanish for an exponential family

$$p_X(x) = \exp\left(\sum_{i=1}^{n} X^i f_i(x) + f_0(x) - \nu(X)\right),$$

hence the name. Here $f_i(x)$ are any functions and $\nu(X)$ the normalizer. We also consider the linear combinations of the e-connection and the metric connection. For example, the m(ixture)-connection $\Gamma^*_{jk,i} = \partial_j g_{ki} - \Gamma_{ij,k}$ satisfies $\Gamma^*_{ij,k} = 0$ for any mixture $p_X(x) = \sum_{l=0}^{n} X^l q_l(x)$ of $n+1$ probability densities $q_l(x)$, where $X^1, \dots, X^n > 0$ and $X^0 = 1 - X^1 - \dots - X^n > 0$. See the book [1] for more on information geometry.

Hereafter we restrict ourselves to the case of the normal distributions, namely we put

$$p_{(m,s)}(x) = \frac{1}{\sqrt{2\pi}s}\exp\left(-\frac{(x-m)^2}{2s^2}\right)$$

for $(m,s) \in \mathbb{H} = \mathbb{R} \times \mathbb{R}_{>0}$. Then the Kullback-Leibler divergence is expressed as

$$\begin{aligned} &D((m,s),(m',s')) \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}s}\exp\left(-\frac{(x-m)^2}{2s^2}\right)\left\{-\frac{(x-m)^2}{2s^2} + \frac{(x-m')^2}{2(s')^2} - \log\frac{s}{s'}\right\}dx \\ &= -\frac{1}{2} + \frac{s^2}{2(s')^2} + 0 + \frac{(m'-m)^2}{2(s')^2} - \log\frac{s}{s'}. \end{aligned}$$

It is easy to read the Fisher metric g and the e-connection ∇ from the cubic approximation

$$D((m+\Delta m, s+\Delta s),(m,s)) \approx \frac{1}{2}\frac{\Delta m^2 + 2\Delta s^2}{s^2} - \frac{1}{3}\frac{\Delta s^3}{s^3}.$$

Namely,

$$g = \frac{dm^2 + 2ds^2}{s^2}, \quad \Gamma_{12,1} = \Gamma_{21,1} = \frac{-2}{s^3}, \quad \Gamma_{22,2} = \frac{-6}{s^3},$$

and the other coefficients $\Gamma_{ij,k}$ vanish. Note that the normal distributions form an exponential family parametrized by $\left(\frac{m}{s^2}, \frac{-1}{2s^2}\right)$. This implies that the geodesics for the e-connection are the horizontal lines, the vertical half-lines, and the upper semi-parabolas. On the other hand, the geodesics for the m-connection are vertical half-lines and upper semi-circles, which coincide with those for the Poincaré metric $\frac{dm^2 + 1ds^2}{s^2} (\neq g)$.

3. The basic model

Let (m, s, M, S) denote the coordinate system of $\mathbb{H} \times \mathbb{H}$. Take the correspondence

$$F: \quad M = -\frac{m}{s^2} \quad \text{and} \quad S = \frac{1}{s}$$
$$\left(\Leftrightarrow \quad \frac{m}{s} + \frac{M}{S} = 0 \quad \text{and} \quad sS = 1\right)$$

and regard it as a submanifold of $\mathbb{H} \times \mathbb{H}$. For any point (m, s, M, S) in $\mathbb{H} \times \mathbb{H}$, we put

$$f(s, m, S, M) = \frac{\left(\frac{M}{S} + \exp(-h)\frac{m}{s}\right)^2 + \exp(-2h) - 1 + 2h}{2},$$

where $h = -\log(sS)$.

Proposition 1. *Suppose that under the correspondence F a point (m', s') of the first factor of $\mathbb{H} \times \mathbb{H}$ is identified with the point (M, S) of the second factor, i.e., $(m', s', M, S) \in F$. Then the function $f(m, s, M, S)$ presents the Kullback-Leibler divergence $D((m, s), (m's'))$.*

Proof.

$$2f\left(s, m, \frac{1}{s'}, \frac{-m'}{(s')^2}\right) = \left(\frac{-m'}{s'} + \frac{s}{s'} \cdot \frac{m}{s}\right)^2 + \frac{s^2}{(s')^2} - 1 - 2\log\frac{s}{s'}$$
$$= 2D((m, s), (m', s'))$$

□

We define the product $(m, s) * (m', s')$ on the first factor of $\mathbb{H} \times \mathbb{H}$ by putting

$$(m, s) * (m's') = \left(m + m', \sqrt{s^2 + (s')^2}\right).$$

It represents the convolution of probability density functions

$$p_{(m,s)*(m',s')}(x) = (p_{(m,s)} * p_{(m,s)})(x).$$

We define another product $(M,S)\cdot(M',S')$ on the second factor by putting

$$(M,S)\cdot(M',S') = \left(\frac{M(S')^2+(M')S^2}{S^2+(S')^2}, \sqrt{\frac{S^2(S')^2}{S^2+(S')^2}}\right).$$

This represents the normalized pointwise product

$$p_{(M,S)\cdot(M',S')} = p_{(M,S)}(x)\cdot p_{(M',S')}(x) \Big/ \int_{-\infty}^{\infty} p_{(M,S)}(x)\cdot p_{(M',S')}(x)dx.$$

We see that a normalized pointwise product presents a Bayesian learning as follows. Suppose that we have a prior probability density $p(x)$ and a likelihood $q(x)$ of a new data. The likelihood $q(x)$ is the conditional probability density of the data under $p(x)$. Then we obtain the posterior probability density $p(x)q(x)/r$ from Bayes' rule, where r is the probability of the data. Since r does not depend on x, we see that it is the normalizer $\int_{-\infty}^{\infty} p(x)q(x)dx$.

Proposition 2. *The above correspondence F on $\mathbb{H}\times\mathbb{H}$ identifies the product $*$ on the first factor (resp. the second factor) with the product $\cdot$ on the second factor (resp. the first factor).*

Proof. Let (m,s,M,S) and (m',s',M',S') be two points on F. Then we have

$$\sqrt{\frac{S^2(S')^2}{S^2+(S')^2}} = \frac{1}{\sqrt{s^2+(s')^2}}$$

and

$$\begin{aligned}\frac{M(S')^2+(M')S^2}{S^2+(S')^2} &= \frac{-\dfrac{m}{s^2}\dfrac{1}{(s')^2} - \dfrac{m'}{(s')^2}\dfrac{1}{s^2}}{\dfrac{1}{s^2}+\dfrac{1}{(s')^2}} \\ &= \frac{-(m+m')}{s^2+(s')^2}.\end{aligned}$$

Further F is invariant under the interchange of the order of the product $\mathbb{H}\times\mathbb{H}$. □

Thus the correspondence F defines a Fourier-like transformation on $\mathbb{H}$, hence the notation.

4. Contact/foliation topology of Sol^3-manifolds

We recall contact/foliation topology of certain T^2-bundles. First we should notice that contact structure is soft enough to interest topologists since Gray's stability [3] says that any smooth homotopy of a contact structure on a closed (i.e., compact borderless) manifold can be realized as an isotopy of the manifold. Further the symplectization of a contact structure is well-defined up to isotopy which preserves the $\mathbb{R}$-fibers. Note also that many foliators imagines contact topology as a "discretization" of foliation theory at least in 3-dimensional case. Honda [5] completed the isotopy classification of contact structures on T^2-bundles essentially in this spirit (precisely, by using convex surfaces instead of holonomy of leaves). Let $A \in SL(2,\mathbb{Z})$ be the monodromy map of a T^2-bundle. In the case where $\mathrm{tr}A > 2$, the T^2-bundle possesses a canonical pair (ξ_+, ξ_-) of positive and negative contact structures associated to the suspension Anosov flow. Several authors studied the isotopy classes of these contact structures in terms of Honda's classification. Kasuya [6], without appealing to the classification, related the contact structures $\xi_\pm$ to Hirzebruch's construction [4] of a cusp. Hereafter we use relevant part of his description of the contact structures $\xi_\pm$. We notice that he also related it to Ryo Furukawa's observation concerning the toric construction of contact submanifold of S^5 in [9].

It is well-known that any element A of $SL(2,\mathbb{Z})$ with $\mathrm{tr}A > 2$ is conjugate to a positive word of $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. Thus we can take a map

$$\mathbb{Z}/r\mathbb{Z} \ni k \mapsto b_k \in \mathbb{Z}_{\geq 2}$$

not identically 2 such that A is conjugate to

$$\begin{bmatrix} s_k & t_k \\ u_k & v_k \end{bmatrix}^{-1} = \left(\begin{bmatrix} b_{k+r-1} & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} b_{k+r-2} & 1 \\ -1 & 0 \end{bmatrix} \cdots \begin{bmatrix} b_k & 1 \\ -1 & 0 \end{bmatrix} \right)^{-1}.$$

Indeed we have

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1} \left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{b_k - 2} \right) \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} b_k & 1 \\ -1 & 0 \end{bmatrix}.$$

Of course, one may start with a given recurring sequence $b_k \geq 2$ since the trace of the above composition map is greater than 2 unless identically $b_k = 2$. The continued fraction

$$w_k = b_k - \cfrac{1}{b_{k+1} - \cfrac{1}{b_{k+2} + \cdots}}$$

is a quadratic irrational number satisfying

$$t_k {w_k}^2 - (s_k - v_k) w_k - u_k = 0.$$

We have $0 < \overline{w}_k < 1 < w_k$, where $\overline{w}_k$ denotes the other solution of the quadratic equation (i.e., the irrational conjugate of ω). Put $c = \log(w_1 \cdots w_r)$. Take any (large) positive constant $K > 0$ to fix the vectors

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ w_1 & -\overline{w}_1 \end{bmatrix}^{-1} \begin{bmatrix} K \\ 0 \end{bmatrix}$$

and

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ w_1 & -\overline{w}_1 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ K \end{bmatrix}.$$

Then we define the quotient manifold $M(x,y,c)=\mathbb{H}\times\mathbb{H}/\sim$ by the equivalences

$$(m,s,M,S)\sim(e^c m, e^c s, e^{-c}s, e^{-c}S),$$
$$(m,s,M,S)\sim(m+x_1, s, M+x_2, S),$$

and

$$(m,s,M,S)\sim(m+y_1, s, M+y_2, S)$$

for any $(m,s,M,S)\in\mathbb{H}\times\mathbb{H}$. Note that the map

$$(m,s,M,S)\mapsto(e^c m, e^c s, e^{-c}s, e^{-c}S)$$

preserves the lattice on the mM-plane generated by $x=(x_1,x_2)$ and $y=(y_1,y_2)$. Thus $M(x,y,c)$ is a 4-manifold diffeomorphic to a T^2-bundle over the open annulus $S^1\times\mathbb{R}$.

Precisely, we take the function $h=-\log(sS)$ on $\mathbb{H}\times\mathbb{H}\subset\mathbb{C}\times\mathbb{C}$. It is strictly plurisubharmonic since the Hessian diag $\left(\frac{1}{4s^2},\frac{1}{4S^2}\right)$ is clearly positive definite. Now we shift our ground to consider, instead of the complex structure J_{std}, the exact symplectic structure $d\lambda_+$ with fixed primitive

$$\lambda_+=-dJ^*_{\mathrm{std}}dh=\frac{dm}{s}+\frac{dM}{S}=\exp\left(\frac{h+t}{2}\right)dm+\exp\left(\frac{h-t}{2}\right)dM,$$

where $t=\log\frac{S}{s}$. The above strong pseudo-convexity is expressed as $dh(X)>0$ for the vector field X which is defined by $\iota_X d\lambda_+=\lambda_+$. The flow foliation of X carries the holonomy invariant transverse contact structure which is defined by the 1-form λ_+. Conversely the symplectic manifold $\mathbb{H}\times\mathbb{H}$ is the symplectization of the section $\{h=0\}$ whose $\mathbb{R}$-fibers are the flow lines of X. The function h descends to the quotient $M(x,y,c)$, so that the 0-level set is a closed contact 3-manifold $M_0(x,y,c)$. Then $M(x,y,c)$ is the symplectization of $M_0(x,y,c)$. The function t also descends to $M(x,y,c)$, so that its restriction to the section $M_0(x,y,c)$ defines a T^2-bundle projection to the circle $\mathbb{R}/2c\mathbb{Z}$. The T^2-fiber is the quotient of the mM-plane by the lattice generated by x and y. The monodromy map $\begin{bmatrix} e^c & 0\\ 0 & e^{-c}\end{bmatrix}$ with respect to the fundamental basis of the mM-plane can also be written as $\begin{bmatrix} s_1 & t_1\\ u_1 & v_1\end{bmatrix}^{-1}\in SL(2,\mathbb{Z})$ with respect to the T^2-basis $\{x,y\}$. This finally leads us to shift our ground to consider the contact structure ξ_+ on the T^2-bundle with monodromy A.

We can take the global frame (e_1,e_2,e_3) of $TM_0(x,y,c)$ by putting

$$e_1=e^{-t/2}\partial_m,\quad e_2=e^{t/2}\partial_M,\quad\text{and}\quad e_3=2\partial_t.$$

It satisfies the sol^3-relations

$$[e_1,e_2]=0,\quad [e_3,e_1]=-e_1,\quad\text{and}\quad [e_3,e_2]=e_2.$$

The dual coframe (e_1^*,e_2^*,e_3^*) satisfies the corresponding relations

$$de_3^*=0,\quad de_1^*=-e_1^*\wedge e_3,\quad\text{and}\quad de_2^*=e_2^*\wedge e_3^*,$$

where

$$e_1^* = e^{t/2}dm, \quad e_2^* = e^{-t/2}dM, \quad \text{and} \quad e_3^* = dt/2.$$

We call $M_0(x,y,c)$ the Sol^3-manifold associated to A. The suspension Anosov flow is the flow generated by e_3. Since it contracts e_1 and expands e_2, the 1-forms $\alpha_\pm = e_1^* \pm e_2^*$ define a pair of positive and negative contact structures $\xi_\pm$, which we call the bi-contact structure associated to the flow. (This ξ_+ coincides with the above one, hence we use the same notation.) The suspension Anosov flow drifts the positive and negative contact structures toward the stable Anosov foliation defined by the equation $e_1^* = 0$. Further we have the negative symplectic form $d\lambda_-$ on $M(x,y,c)$ which is the symplectization of the negative contact structure ξ_-, where

$$\lambda_- = e^{h/2}(e_1^* - e_2^*) = \exp\left(\frac{h+t}{2}\right)dm - \exp\left(\frac{h-t}{2}\right)dM$$

We have the following contact 5-manifold which is crucial in contemporary contact topology since it shows how to generalize the classical propeller construction of 3-dimensional contact structure to higher dimensions so that the propeller keeps its relation to convex (hyper)surface theory especially to Giroux torsion (see [10] and [7]). Take $\eta = e_1^* + \cos\theta e_2^* + \sin\theta d\varphi$ on $M_0(x,y,c)\times S^1\times\mathbb{R}$, where $\theta \in S^1$ and $\varphi \in \mathbb{R}$ are the coordinates. It defines a positive contact structure since

$$\begin{aligned}\eta\wedge(d\eta)^2 &= (e_1^* + \cos\theta e_2^* + \sin\theta d\varphi)\\ &\qquad \wedge(e_3^*\wedge e_1^* - \sin\theta d\theta\wedge e_2^* + \cos\theta e_2^*\wedge e_3^* + \cos\theta d\theta\wedge d\varphi)^2\\ &= (4\cos^4\theta + 2\sin^2\theta)e_1^*\wedge e_2^*\wedge e_3^*\wedge d\theta\wedge d\varphi > 0.\end{aligned}$$

We divide the hypersurface $\{\varphi = 0\}$ into two regions $\Sigma^\pm$ according to the sign of $\sin\theta$. Then the restrictions of the 2-form $d\eta$ on $\pm\mathrm{int}\Sigma^\pm$ are symplectic structures, where $-\mathrm{int}\Sigma^-$ denotes $\mathrm{int}\Sigma^-$ with reversed orientation. We can see that $\pm\mathrm{int}\Sigma^\pm$ are symplectomorphic to each other by switching the sign of θ-coordinate. We further divide the region Σ^+ into two regions Σ^+_+ and Σ^+_- according to the sign of $\cos\theta$. The following observation fits the scope of [8].

PROPOSITION 3. $\mathrm{int}\Sigma^+_+\times\mathbb{R}$ *(resp.* $\mathrm{int}\Sigma^+_-\times\mathbb{R}$*) is contactomorphic to* $\ker(\lambda_+ + d\varphi)$ *(resp.* $\ker(\lambda_- + d\varphi)$*) on* $M(x,y,c)\times\mathbb{R}$ *(resp.* $-M(x,y,c)\times\mathbb{R}$*).*

PROOF. We have

$$\frac{\eta}{\sin\theta} = e^{(H+T)/2}dm \pm e^{(H-T)/2}dM + d\varphi,$$

where $H = \log\dfrac{|\cos\theta|}{\sin^2\theta}$ and $T = t - \log|\cos\theta|$. □

Namely, we consider the symplectic manifold $\mathrm{int}\Sigma^+$ as the result of the next procedure expressing that scope. The 4-manifold $M(x,y,c)$ with the conformal class of the positive symplectic structure $d\lambda_+$ and its orientation-reversion $-M(x,y,c)$ with the conformal class of the *positive* symplectic structure $d\lambda_-$ are pasted together so as to be realized as the symplectic region intΣ^+ along the Levi-flat ends each of which is expressed as $\{h = +\infty\} \approx M_0(x, y, c)$ and carries the stable Anosov foliation $e_1^* = 0$. This can be viewed as the interchange of the sign of the bi-contact structure at the limit of the Anosov flow via the stable foliation. That is why we consider the positive symplectic structure $d\lambda_+$ as paired with the negative one $d\lambda_-$ on $M(x, y, c)$. (This idea also comes from the topology of Hirzebruch-Inoue or hyperbolic Inoue surfaces.)

5. Results

Now we investigate the model in §3 in the light of the description in §4.

PROPOSITION 4. *The correspondence F and the function f are invariant under the monodromy map $(m, s, M, S) \mapsto (e^c m, e^c s, e^{-c} s, e^{-c} S)$.*

PROOF. The monodromy map preserves the inverse coefficients of variance $\frac{m}{s}$ and $\frac{M}{S}$, and the strictly plurisubharmonic function $h = -\log(sS)$. □

PROPOSITION 5. *The symplectic structure*

$$d\lambda_+ = \frac{dm \wedge ds}{s^2} + \frac{dM \wedge dS}{S^2},$$

the sum of Fisher metrics

$$g = \frac{dm^2 + 2ds^2}{s^2} + \frac{dM^2 + 2dS^2}{S^2},$$

and the almost complex structure

$$J : \partial_m \mapsto \frac{1}{\sqrt{2}}\partial_s, \partial_M \mapsto \frac{1}{\sqrt{2}}\partial_S$$

on $\mathbb{H} \times \mathbb{H}(\ni (m, s, M, S))$ satisfy $g(\cdot, \cdot) = \sqrt{2}d\lambda_+(\cdot, J\cdot)$ and descend to the quotient $M(x, y, c)$.

PROOF. We have

$$\sqrt{2}d\lambda_+(\partial_m, \frac{1}{\sqrt{2}}\partial_s) = \frac{1}{s^2}$$

and

$$\sqrt{2}d\lambda_+(\partial_s, -\sqrt{2}\partial_m) = \frac{2}{s^2}.$$

The rest is clear. □

PROPOSITION 6. *The "Fourier" correspondence $F \subset (\mathbb{H} \times \mathbb{H}, d\lambda_+)$ is a smooth surface contained in the contact-type hypersurface $H = \{sS = 1\}$ with positive contact structure* $\ker \alpha_+$. *It is the union of Legendrian lines*

$$\{e^{T/2}m + e^{-T/2}M = 0 \quad \textit{and} \quad t = T\} \quad (T \in \mathbb{R}).$$

These lines descend to the quotient $M_0(x, y, c)$ of H as Legendrian curves which form a dense immersion of the surface F.

PROOF. We see that

$$e^{t/2}m + e^{-t/2}M = 0 \quad \Leftrightarrow \quad \frac{m}{s} + \frac{M}{S} = 0.$$

Together with $sS = 1$, this defines F. The Legendrian lines for $T = T_0 + 2c$ and $T = T_0$ descend to the same Legendrian immersed curve on $(M_0(x, y, c), \ker \alpha_+)$, which is either closed or dense in the toral fiber $\{t = T\} \subset M_0(x, y, c)$ depending on whether the slope $-e^T$ is rational or irrational with respect to the basis $\{x, y\}$. □

PROPOSITION 7. *We have a negative symplectic structure*

$$d\lambda_- = \frac{dm \wedge ds}{s^2} - \frac{dM \wedge dS}{S^2}$$

on $\mathbb{H} \times \mathbb{H}$, with respect to which the surface F is a Lagrangian correspondence.

PROOF. The tangent space of F is expressed as

$$TF = \langle s\partial_m - S\partial_M, m\partial_m + s\partial_s - M\partial_M - S\partial_S \rangle.$$

Then we have

$$\frac{dm \wedge ds(s\partial_m, m\partial_m + s\partial_s)}{s^2} - \frac{dM \wedge dS(-S\partial_M, -M\partial_M - S\partial_S)}{S^2} = 1 - 1 = 0.$$

□

The correspondence F also satisfies the next interesting property.

PROPOSITION 8. *We call a geodesic for the e-connection an e-geodesic. Any e-geodesic on the first factor of $\mathbb{H} \times \mathbb{H}$ corresponds to an e-geodesic on the second factor via F, and vice-versa.*

PROOF. We have

$$S(k_1 s^2 + k_2 m + k_3) = s(k_3 S^2 + k_2 M + k_1) = 0,$$

where k_i are any constants. □

The next theorem shows that a certain contact flow describes Bayesian learning processes.

THEOREM 1. *The contact Hamiltonian vector field Y of the restriction of the inverse coefficient of variation $\frac{m}{s}$ to the hypersurface $\{sS = 1\} \subset \mathbb{H} \times \mathbb{H}$ with respect to α_+ is expressed as*

$$Y = \frac{m}{s}e_1 - \frac{1}{2}e_3 = m\partial_m - \frac{1}{2}(S\partial_S - s\partial_s)|_{\{sS=1\}}.$$

It is tangent to the surface F. Let Y_i denote the push-forward of the restriction $Y|_F$ to the i-th factor of $\mathbb{H} \times \mathbb{H}$ ($i = 1, 2$). Then the flow of Y_1 comes out of the origin $(0, 0) \in \overline{\mathbb{H}}$ along the half-parabolic (or half-liner) e-geodesics, and the flow of Y_2 goes into the M-axis $\{S = 0\} \subset \overline{\mathbb{H}}$ along the vertical half-liner e-geodesics. The former can be discretized into the iterations of convolutions and therefore the latter into those of Bayesian learnings.

Proof. The interior product $\iota_Y d\left(\frac{m}{s}+\frac{M}{S}\right) = \frac{1}{2}\left(\frac{m}{s}+\frac{M}{S}\right)$ vanishes along F. The Lie derivative of α_+ is $\mathcal{L}_Y\alpha_+ = \frac{1}{2}\alpha_+$. The Hamiltonian function is $\iota_Y\alpha_+ = \frac{m}{s}$. The rest is easy. □

Remark 1. *Let α be a positive contact form on an oriented 3-manifold M_0, i.e., a 1-form satisfying $\alpha \wedge d\alpha > 0$. Then the 2-form $d(e^h\alpha)$ on the product $\mathbb{R}\times M_0 = M$ ($h \in \mathbb{R}$) defines the symplectic structure as the symplectization of the contact structure. Given a function H on M_0, we can define the contact Hamiltonian vector field Y for H as the well-defined push-forward of the usual Hamiltonian vector field $\widetilde{Y}$ for the function $\widetilde{H} = e^h H$ to M_0. Then we have $\alpha(Y) = H$. Note that Y is contact, i.e., it preserves the contact structure* $\ker\alpha$ *since $\mathcal{L}_{\widetilde{Y}} e^h\alpha = \iota_{\widetilde{Y}} d(e^h\alpha) + d\iota_{\widetilde{Y}}(e^h\alpha) = -d(e^h H) + d(e^h H) = 0$. From the non-integrability of the contact structure* $\ker\alpha$*, we see that a contact vector field Y' with $\alpha(Y') = 0$ must be zero. Thus any contact vector field is a contact Hamiltonian vector field. If we fix $\widetilde{H}$ and change the section of the $\mathbb{R}$-fibration of M, we obtain another pair of contact form $e^f\alpha$ and contact Hamiltonian function $e^{-f}H$ on the base manifold M_0 which defines the same contact Hamiltonian vector field Y. This means that, for a fixed contact structure, each contact form presents an isomorphism between the space of functions and the space of contact vector fields as well as a section of the symplectization.*

6. Problems

Since information geometry concerns parameter spaces with geometric structures, it would have some relation to moduli theory. This was one of the starting points of this research. In the present, we have no intrinsic relation between the pairs of normal distributions and abelian surfaces.

Problem 1. Is there any number theoretical relation between two normal distributions which enables us to relate the pair to an abelian surface with real multiplication ?

The other starting point was the following splitting result proved in [8] (see also [2]): A non-singular flow on a closed 3-manifold is projectively Anosov if and only if it is simultaneously tangent to a mutually transverse pair of positive and negative contact structures, i.e., to a bi-contact structure. In order to consider the topology of the mixed parameter space, as a closed manifold or as a space with non-trivial topology, we have to know how to split it into individual parameter spaces. This motivates us to consider Anosov flows. However, this kind of hyperbolicity usually makes the dynamics chaotic under the compactness of the space (like Arnold's cat map). Thus it is reasonable to consider quotient manifolds of an already split space. The author is trying to find another example than the pair of normal distributions from this point of view.

There are still questionable points on our own model in this article. For example,

Problem 2. Is there any geometrical reason for taking the particular function f ?

The author consider them as future tasks and raises the above one as an open problem.

References

[1] Shun-ichi Amari, Information geometry and its applications, App. Math. Sci. , Springer, Tokyo, 2016.

[2] Yakov Eliashberg and William Thurston, Confoliations, A. M. S. University Lecture Series **13** (1997).

[3] John Gray, Some global properties of contact structures, Ann. of Math. **69(2)** (1959), 421450

[4] Friedrich Hirzebruch, Hilbert modular surfaces, L'Ensignement Math. **19** (1973), 183-28.

[5] Ko Honda, On the classification of tight contact structures II, J. Diff. Geom. **55** (2000), 83-143.

[6] Naohiko Kasuya, The canonical contact structure on the link of a cusp singularity, Tokyo J. Math. **37** (2014), 1-20.

[7] Patrick Massot, Klaus Niederkrüger, and Chris Wendl, Weak and strong fillability of higher dimensional contact manifolds, Invent. Math. **192**(2)(2013), 287-373.

[8] Yoshihiko Mitsumatsu, Anosov flows and non-Stein symplectic manifolds, Ann. Inst. Fourier **45** (1995), 1407-1421.

[9] Atsuhide Mori, The Reeb foliation arises as a family of Legendrian submanifolds at the end of a deformation of the standard S^3 in S^5, C. R. Acad. Sci. Paris, Ser. I **350** (2012), 67-70.

[10] Atsuhide Mori, On higher dimensional violation of Thurston-Bennequin inequality, preprint (2015). This includes "Reeb foliations on S^5 and contact 5-manifolds violating the Thurston-Bennequin inequality" (arXiv:0906.3237) and "On the violation of Thurston-Bennequin inequality for a certain non-convex hypersurface" (arXiv:1111.0383).