Uhlmann's parallelism

and

Nagaoka's quantum information geometry

Keiji Matsumoto

METR 97-09 October 1997

Uhmann's parallelism and Nagaoka's quantum information geometry

Keiji Matsumoto [1]

Abstract

In the space of density matrix, two estimation-theoretically natural geometrical structures can be introduced: Uhlmann's parallelism and Nagaoka's quantum information geometry. In this paper, intrinsic relation between them is clearified.

Keywords: Uhlmann's curvature, fiber bundle, information geometry, e-connection, quantum estimation theory, Cramer-Rao type bound

[1]Department of Mathematical Engineering and Information Physics
University of Tokyo, Bunkyo-ku,Tokyo 113, Japan

1 Introduction

Quantum estimation theory deals with the identification of the density matrix or given system by use of data produced by appropriately desigened experiment [4][5][6]. From this quantum-estimation-theoretical point of view, there are two natural geometridal structures in the space of full rank density matrices. One is Uhlmann's parallelism and the other is Nagaoka's information geometry.

Uhlmann's parallelism is generalization of Berry's phase, which, by far confirmed by several experiments, is a holonomy of a natural connection in the line bundle over the space of pure states [1][3][12][13]. In 1986, Uhlmann generalized the theory to include mixed states in the Hilbert space $\mathcal{H}$ [14][15] [16].

It is pointed out that the Uhlmann's parallelism is deeply concerned with quantum estimation theory. Concretely speaking, iff Uhlmann's curvature vanishes, SLD CR bound is attained, which implies the non-commutative nature of quantum mechanics is not relevant in the model [8] .

In the classical estimation theory (in this paper, the estimation theory of probability distributions is called classical estimation theory in the sense the theory is not quantum nechabical), S. Amari and their collaborators have shown that a 'dual connection' plays important role in the higher-order asymptotic theory. This geometrical theory of statistics is known as 'information geometry' [2]. Nagaoka, one of Amari's collaborators, formulated the quantum version of information geometry and showed that the geometry nicely characterizes models which have efficient estimators.

This paper sheds light on the intrinsic relation between two geometries. Also, a kind of 'duality' between SLD and RLD, both of which are important concept in the quantum estimation theory, is pointed out, and applied to the consideration of the canonical distribution.

2 Horizontal lift and SLD

In this paper, $d \equiv \dim \mathcal{H}$ is assumed to be finite for the sake of clarity, and desnsity matrices are full rank. The author believe the essence of the discussion will not be damaged by this restriction.

Let $\mathcal{W}_d$ be the space of $d \times d$ complex and full-rank matrix W such that

$$\mathrm{tr} W W^* = 1,$$

$\mathcal{P}_+$ the space of density operators whose rank is d, and π the map from $\mathcal{W}_d$ to $\mathcal{P}_+$ such that

$$\rho = \pi(W) \equiv W W^*.$$

Because $\pi(WU)$ is identical to $\pi(W)$ iff U is a $d \times d$ unitary matrix, it is natural to see the space $\mathcal{W}_d$ as the total space of the *principal fiber bundle* with the *base space* $\mathcal{P}_+$ and the *structuregroup* $U(d)$ [7]. One possible physical interpretation of W is a representation of a state vector $|W\rangle$ in a bigger Hilbert space $\mathcal{H} \otimes \mathcal{H}'$. Here, the dimension of $\mathcal{H}'$ is r and the operation $\pi(*)$ corresponds to the partial trace of $|W\rangle\langle W|$ over $\mathcal{H}'$.

In this section, basic concepts about the tangent bundle $\mathcal{T}(\mathcal{W}_d)$ over $\mathcal{W}_d$, which is a real manifold with the real parameter $\zeta = (\zeta^1, ..., \zeta^{2d^2-1})^T$, are introduced.

The *matrix representation* $\mathbf{M}(\partial/\partial\zeta^i)$ of the tangent vector $\partial/\partial\zeta^i$ (throughout the thesis, the tangent vector is understood as the differential operator) is a $d \times r$ complex matrix such that

$$\mathbf{M}\left(\frac{\partial}{\partial \zeta^i}\right) \equiv 2 \frac{\partial}{\partial \zeta^i} W(\zeta).$$

The real span of the matrix representations is

$$\{X \mid \mathrm{Re}\,\mathrm{tr} X W^*(\zeta) = 0, X \in M(d, r, \mathbf{C})\}.$$

We introduce the inner product $\langle\langle *, * \rangle\rangle_W$ to $\mathcal{T}(\mathcal{W}_d)$ such that,

$$\begin{aligned} & \langle\langle \hat{X}, \hat{Y} \rangle\rangle_W \\ \equiv \; & \sum_{i,j} (\mathrm{Re}(\mathbf{M}\hat{X})_{ij} \mathrm{Re}(\mathbf{M}\hat{Y})_{ij} + \mathrm{Im}(\mathbf{M}\hat{X})_{ij} \mathrm{Im}(\mathbf{M}\hat{X})_{ij}) \\ = \; & \mathrm{Re}\,\mathrm{tr}(\,(\mathbf{M}\hat{X})(\mathbf{M}\hat{Y})^*\,), \end{aligned}$$

which is invariant under the action of $U \in U(d)$ to the matrix representation of the tangent vector from right side,

$$U \in U(d),$$
$$\langle\langle\, (\mathbf{M}\hat{X})U, (\mathbf{M}\hat{Y})U \rangle\rangle_{WU} = \langle\langle \mathbf{M}\hat{X}, \mathbf{M}\hat{Y} \rangle\rangle_{W}$$

Let us decompose $\mathcal{T}_W(\mathcal{W}_d)$ into the direct sum of the *horizontal subspace* $\mathcal{LS}_W$ and the *vertical subspace* $\mathcal{K}_W$ where $\mathcal{LS}_W$ is defined by

$$\mathcal{LS}_W \equiv \{\hat{X} \mid W^*(\mathbf{M}\hat{X}) = (\mathbf{M}\hat{X})^* W\}, \tag{1}$$

and $\mathcal{K}_W$ is the orthogonal complement space $\mathcal{T}_W(\mathcal{W}_d) \ominus \mathcal{LS}_W$ with respect to the inner product $\langle\langle *, * \rangle\rangle_W$. $\hat{X} \in \mathcal{K}_W$ satisfies

$$(\mathbf{M}\hat{X})W^* + W(\mathbf{M}\hat{X})^* = 0, \tag{2}$$

or its equivalence,

$$\pi_*(\hat{X}) = 0, \tag{3}$$

where π_* is the differential map of π. A member of the horizontal subspace and the vertical subspace are called a *horizontal vector* and *vertical vector*, respectively. The image of $\hat{X} \in \mathcal{T}_W(\mathcal{W}_d)$ by the projection onto the horizontal subspace $\mathcal{LS}_W$ is called the *horizontal component*, while the image by the projection onto the vertical subspace $\mathcal{K}_W$ is called *vertical component*.

The *horizontal lift* h_W is a mapping from $\mathcal{T}_{\pi(W)}(\mathcal{P}_+)$ to $\mathcal{T}_W(\mathcal{W}_d)$ such that

$$\pi_*\left(\underset{W}{h}(X)\right) = X,$$
$$\underset{W}{h}(X) \in \mathcal{LS}_W.$$

Because of the following theorem, the matrix representation of the horizontal lift $\pi_*(h_W(X))$ is a representation of the tangent vector $X \in \mathcal{T}_{\pi(W)}(\mathcal{P}_+)$.

Theorem 1 *h_W is a isomorphism from $\mathcal{T}_{\pi(W)}(\mathcal{P}_+)$ to $\mathcal{LS}_W$.*

Proof First, notice that for any $\hat{Y} \in \mathcal{T}_W(\mathcal{W}_d)$, $W + \varepsilon\mathbf{M}(\hat{Y})$ also is a member of $\mathcal{W}_d$ if ε is small enough. Therefore, we have

$$\pi_*(\,\mathcal{T}_W(\mathcal{W}_d)\,) \subset \mathcal{T}_{\pi(W)}(\mathcal{P}_+).$$

Second, we prove that the map $\pi_*|_{\mathcal{LS}_W}$ is a one to one map from $\mathcal{LS}_W$ to $\mathcal{T}_{\pi(W)}(\mathcal{P}_+)$. For that, it is sufficient to prove that $\hat{X} = 0$ when $\hat{X} \in \mathcal{K}_W$. This statement is proved to be true because $\mathcal{K}_W$ is orthogonal to $\mathcal{LS}_W$.

Finally, checking the dimension of $\mathcal{T}_{\pi(W)}(\mathcal{P}_+)$ is equal to $\mathcal{LS}_W$, we have the theorem. □

Using the horizontal lift, the inner product $\langle *, * \rangle$ in $\mathcal{T}(\mathcal{P}_+)$ is deduced from $\langle\langle *, * \rangle\rangle$:

$$\langle X, Y \rangle_{\pi(W)} = \left\langle\left\langle \underset{W}{h}(X), \underset{W}{h}(Y) \right\rangle\right\rangle_W .$$

The horizontal lift h satisfies the following equality so that the above definition of the inner product $\langle *, * \rangle$ is self-consistent:

$$\left\langle\left\langle \underset{W}{h} X, \underset{W}{h} Y \right\rangle\right\rangle = \left\langle\left\langle \underset{WU}{h} X, \underset{WU}{h} Y \right\rangle\right\rangle, \ (U, U' \in U(n)\ ).$$

The *symmetrized logarithmic derivative* (SLD, in short) of $X \in \mathcal{T}_{\pi(W)}(\mathcal{P}_+)$ is the Hermitian operator L^S_X in $\mathcal{H}$ defined by the equation

$$X\rho(\theta) = \frac{1}{2}(L^S_X \rho(\theta) + \rho(\theta) L^S_X), \tag{4}$$

where θ is a real parameter which is assigned to a member of $\mathcal{P}_+$. Iff the density operator is strictly positive, SLD is uniquely defined by (4). $L^S_{\partial/\partial\theta^i}$ is often denoted simply by L^S_i.

SLD is closely related to the horizontal lift by the following equation:

$$\mathbf{M}\left(\underset{W}{h} X\right) = L^S_X W. \tag{5}$$

3 Definition of Uhlmann's parallelism

Berry's phase, by far confirmed by several experiments, is a holonomy of a natural connection in the line bundle over the space of pure states [1][3]. In 1986, Uhlmann generalized the theory to include mixed states in the Hilbert space $\mathcal{H}$ [14][15] [16]. For notational simplicity, the argument θ is omitted, as long as the omission is not misleading.

Define a *horizontal lift* of a curve $C = \{\rho(t) | t \in \mathbf{R}\}$ in $\mathcal{P}_+$ as a curve $C_h = \{W(t) | t \in \mathbf{R}\}$ in $\mathcal{W}_d$ which satisfies $C = \pi(C_h)$ and

$$\frac{dW(t)}{dt} = \mathbf{M}\left(\underset{W(t)}{h} \left(\frac{d}{dt} \right) \right). \tag{6}$$

Then, the *relative phase factor* (RPF) between $\rho(t_0)$ and $\rho(t_1)$ along the curve C is the unitary matrix U which satisfies the equation

$$W(t_1) = \hat{W}_1 U,$$

where $\hat{W}_1$ satisfies $\rho(t_1) = \pi(\hat{W}_1)$ and

$$\hat{W}_1^* W(t_0) = W^*(t_0) \hat{W}_1.$$

RPF is said to vanish when it is equal to the identity.

4 RPF for infinitesimal loop

The RPF for the infinitesimal loop

$$\begin{array}{ccc} (\theta^1, ..., \theta^i, ..., \theta^j + d\theta^j, ..., \theta^m) & \leftarrow & (\theta^1, ..., \theta^i + d\theta^i, ..., \theta^j + d\theta^j, ..., \theta^m) \\ \downarrow & & \uparrow \\ \theta = (\theta^1, ..., \theta^i, ..., \theta^j, ..., \theta^m) & \rightarrow & (\theta^1, ..., \theta^i + d\theta^i, ..., \theta^j, ..., \theta^m) \end{array} \tag{7}$$

is calculated up to the second order of $d\theta$ by expanding the solution of the equation (6) to that order:

$$\begin{aligned} & I + \frac{1}{2} W^{-1} F_{ij} W \; d\theta^i d\theta^j + o(d\theta)^2, \\ & F_{ij} = (\partial_i L_j^S - \partial_j L_i^S) - \frac{1}{2}[L_i^S, L_j^S]. \end{aligned} \tag{8}$$

Note that F_{ij} is a representation of the curvature form, and that RPF for any closed loop vanishes iff F_{ij} is zero at any point in $\mathcal{M}$.

5 Nagaoka's quantum information geometry

In this section, we give brief review of Nagaoka's quantum information geometry, which is another geometrical theory of the quantum statistical model than Uhlmann's parallelism.

In the Nagaoka's geometry, metric tensor is chosen to be SLD Fisher information matrix. Letting $\mathbf{L}^S$ denote the linear mapping from the tangent vector to its SLD, *e-parallel transport* $\coprod^{(e)}$ is defined as follows:

$$\mathbf{L}^S\left(\coprod_{\rho\to\sigma}^{(e)} X\right) \equiv \mathbf{L}^S X - \mathrm{tr}(\sigma \mathbf{L}^S X). \tag{9}$$

Note that in the faithful model, $\mathbf{L}^S$ is one to one mapping, and the equation (9) defines the connection uniquely. The dual $\coprod^{(m)}$ of e-parallel transport with respect to SLD inner product is called *m-parallel transport*,

$$\forall X, Y \in \mathcal{T}_\rho(\mathcal{P}_+) \ \left\langle \coprod_{\rho\to\sigma}^{(m)} X, \coprod_{\rho\to\sigma}^{(e)} Y \right\rangle_\sigma = \langle X, Y\rangle_\sigma.$$

For the autoparallel manifold in e-connection, the e- covariant derivative is calculated as

$$\mathbf{L}^S\left(\nabla_X^{(e)} Y\right) = X L_Y^S - \mathrm{tr}\rho X L_Y^S,$$

and the tortion of e-connection $T^{(e)}$ as,

$$\begin{aligned} \mathbf{L}^S\left(T^{(e)}(X,Y)\right) &= \mathbf{L}^S\left(\nabla_X^{(e)} Y - \nabla_Y X - [X,Y]\right) \\ &= X\mathbf{L}^S(Y) - Y\mathbf{L}^S(X), \end{aligned} \tag{10}$$

or, equivalently,

$$T^{(e)}(X,Y)\rho(\theta) = \frac{1}{4}[[L_X^S, L_Y^S], \rho], \tag{11}$$

where X and Y are understood as differential operators.

Nagaoka showed that these e- and m-connections nicely characterize the estimation theoretical properties of models, in a different manner than Uhlmann's parallelism. Is there any relation between the two geometrical structures?

6 w-connection in $\mathcal{W}$

To elucidate the relations between Uhlmann's parallelism and Nagaoka's information geometry, we consider the geometry of the tangent bundle over the total space $\mathcal{W}_d$.

the logarithmic derivative $\mathbf{L}(\hat{X})$ of $X \in \mathcal{T}_W(\mathcal{W}_d)$ is a $d \times d$ complex matrix which satisfies the equation,

$$\mathbf{L}(\hat{X})W = \mathbf{M}(\hat{X}). \tag{12}$$

$\mathbf{L}(\partial/\partial\zeta^i)$ is often denoted by $L_{\partial/\partial\zeta^i}$ for simplicity. Notice the logarithmic derivative is uniquely defined.

The real span of the logarithmic derivatives of at ζ is

$$\{L \mid \mathrm{Re}\,\mathrm{tr}(L\pi(W(\zeta))) = 0, L \in M_d(\mathbf{C})\}.$$

We introduce the *w-connection* in $\mathcal{T}(\mathcal{W}_d)$ by the *w-parallel transport* defined by

$$\mathbf{L}\left(\coprod_{W_0 \to W_1}^{(w)} \hat{X}\right) \equiv \mathbf{L}\hat{X} - \mathrm{Re}\,\mathrm{tr}(\pi(W_1)\mathbf{L}\hat{X}). \tag{13}$$

This w-parallel transport is left invariant under the action of $U \in U(d)$ in the following sense:

$$\mathbf{L}\left(\coprod_{W_0 \to W_1}^{(w)} \hat{X}\right) = \mathbf{L}\left(\coprod_{W_0 U \to W_1}^{(w)} \hat{X}\right) = \mathbf{L}\left(\coprod_{W_0 \to W_1 U}^{(w)} \hat{X}\right).$$

The covariant derivative $\nabla^{(w)}$ and the tortion w-tortion $T^{(w)}(\hat{X}, \hat{Y})$ are easily calculated for the w-autoparallel submanifold of $\mathcal{W}_d$ as

$$\mathbf{L}\left(\nabla_{\hat{X}}^{(w)} \hat{Y}\right) \quad = \quad \hat{X}\mathbf{L}(\hat{Y}) - \mathrm{Re}\,\mathrm{tr}\left(\pi(W)\hat{X}\mathbf{L}(\hat{Y})\right) I$$

$$
\begin{aligned}
\mathbf{L}T^{(w)}(\hat{X},\hat{Y}) &= \mathbf{L}\left(\nabla^{(w)}_{\hat{X}}\hat{Y} - \nabla^{(w)}_{\hat{Y}}\hat{X} - [\hat{X},\hat{Y}]\right) \\
&= \frac{1}{2}[L_X, L_Y] - \mathrm{Re}\,\mathrm{tr}\left\{\pi(W)(\hat{X}L_Y - \hat{Y}L_X)\right\} I,
\end{aligned}
\tag{14}
$$

where the tangent vector $\hat{X}$ is understood as a differential operator, and I is the identity in the Hilbert space $\mathcal{H}$. The latter equation is equivalent to

$$
\mathbf{M}T^{(w)}(\hat{X},\hat{Y}) = \frac{1}{2}[L_X, L_Y]W - \mathrm{Re}\,\mathrm{tr}\left\{\pi(W)(\hat{X}L_Y - \hat{Y}L_X)\right\} W, \tag{15}
$$

which is of use when the theory is generalized to non-faithful models.

7 Projection of geometric structures

In the beginning, we show that the Nagaoka's information geometry is naturally induced from the geometry of the $\mathcal{T}(\mathcal{W})$.

As in the definition, the metric $\langle *, * \rangle$ in Nagaoka's quantum information geometry is induced from the natural metric $\langle\langle *, * \rangle\rangle$ in $\mathcal{T}(\mathcal{W})$.

Not only the metric $\langle *, * \rangle$, but also the transport $\coprod^{(e)}$ is induced from $\coprod^{(w)}$:

$$
\begin{array}{ccc}
\rho = \pi(W) & & \sigma = \pi(V) \\
 & \coprod^{(w)} & \\
\hat{X} = \underset{W}{h}(X) \in \mathcal{T}_W(\mathcal{W}) & --\longrightarrow & \coprod^{(w)}_{W\to V}\hat{X} \in \mathcal{T}_V(\mathcal{W}) \\
h\uparrow & & \downarrow \pi_* \\
X \in \mathcal{T}_\rho(\mathcal{P}_+) & --\longrightarrow & \coprod^{(e)}_{\rho\to\sigma} X \in \mathcal{T}_\sigma(\mathcal{P}_+) \\
 & \coprod^{(e)} &
\end{array}
\tag{16}
$$

The horizontal lift h satisfies the following requirements so that the definition of the transport $\coprod^{(e)}$ by the diagrams above are consistent:

$$
\pi_*\left(\coprod^{(w)}_{W\to V}\underset{W}{h}(X)\right) = \pi_*\left(\coprod^{(w)}_{WU\to VU'}\underset{WU}{h}(X)\right),
$$

Because of the diagram, it is quite easy to see that if the submanifold $\mathcal{N}$ of $\mathcal{W}$ is w-autoparallel, the model $\mathcal{M} = \pi(\mathcal{N})$ is e-autoparallel.

Some elementary calculations leads to the following theorem, which illustrates the relation between the two geometries.

Theorem 2 *Let $\mathcal{M}$ be a submanifold of $\mathcal{P}_+$ which is induced from w-autoparallel submanifold $\mathcal{N}$ of $\mathcal{W}$ by $\mathcal{M} = \pi(\mathcal{N})$ and X and Y tangent vectors to $\mathcal{P}_+$ such that,*

$$\begin{aligned} X &= \sum_i x^i \frac{\partial}{\partial \theta^i}, \\ Y &= \sum_i y^i \frac{\partial}{\partial \theta^i}. \end{aligned}$$

$T^{(w)}(h(X), h(Y))$ is decomposed into the sum such that

$$\begin{aligned} \mathbf{L}T^{(w)}(h(X), h(Y)) &= \mathbf{L}^S T^{(e)}(X, Y) - F_{XY}, \\ F_{XY} &= \sum_{i,j} F_{ij} x^i y^j, \end{aligned}$$

where $\mathbf{L}^{-1}(\mathbf{L}^S T^{(e)}(X, Y))$ is a horizontal vector and $\mathbf{L}^{-1} F_{XY}$ is a vertical subspace.

In other words, the horizontal component of the w-tortion is the e-tortion and the vertical component is the curvature form of the Uhlmann parallelism.

8 The duality of SLD and RLD

First, we define the *right logarithmic derivative* (RLD, in short), which played quite important role in the estimation theory of the Gaussian model, which is a superposition of the coherent states by the Gaussian kernel (see Ref. [17] and pp. 80-90 of Ref. [6]).

RLD $L_i^R(\theta)$ of the parameter θ^i is defined by the equation

$$\frac{\partial \rho(\theta)}{\partial \theta^i} = L_i^R(\theta) \rho(\theta),$$

and RLD $\mathbf{L}^R X$ of the tangent vector by the equation

$$\hat{X}\rho(\theta) = (\mathbf{L}^R X)\rho(\theta).$$

Our question is why we need two types of logarithmic derivatives, SLD and RLD namely, and what the relations between them. To answer the question, we interpret the total space $\mathcal{W}_d$, like in the section 3, as the space of the state vector $|\Phi\rangle$ in the bigger Hilbert space $\mathcal{H} \otimes \mathcal{H}'$, where we took the dimension of $\mathcal{H}'$ to be d, the dimension of $\mathcal{H}$. Let us call $\mathcal{H}$ the observed system, and $\mathcal{H}'$ the hidden system, and the partial trace over $\mathcal{H}$ and $\mathcal{H}'$ are denoted by π and π', respectively.

In terms of W, π and π' write

$$\begin{aligned} \pi(W) &= WW^* = \rho, \\ \pi'(W) &= W^*W = \sigma, \end{aligned}$$

and π_* and π'_* write

$$\begin{aligned} \pi_*(\hat{X})\rho &= \frac{1}{2}(\,(\mathbf{M}\hat{X})W^* + W(\mathbf{M}\hat{X})^*\,), \\ \pi'_*(\hat{X})\sigma &= \frac{1}{2}(\,(\mathbf{M}\hat{X})^*W + W^*(\mathbf{M}\hat{X})\,), \end{aligned} \tag{17}$$

The RLD subspace $\mathcal{LR}_W$ of $\mathcal{T}_W(\mathcal{W}_d)$ is the space of all vectors which satisfies,

$$\mathbf{L}(\hat{X}) = \mathbf{L}^R(\pi_*(\hat{X})), \tag{18}$$

or, its equivalence,

$$(\mathbf{M}\hat{X})W^* = W(\mathbf{M}\hat{X})^*. \tag{19}$$

Then, from (17), (19) and (1), we have

$$\begin{aligned} \mathcal{LR}_W &= \{\hat{X} \mid \mathbf{L}(\hat{X}) = \mathbf{L}^S(\pi'_*(\hat{X}))\}, & (20) \\ \mathcal{LS}_W &= \{\hat{X} \mid \mathbf{L}(\hat{X}) = \mathbf{L}^R(\pi'_*(\hat{X}))\}. & (21) \end{aligned}$$

In other words, looking from the hidden system, the RLD subspace looks like the horizontal subspace, and the horizontal subspace looks like the RLD subspace. We call the fact (21) the *duality between SLD and RLD.*

The orthogonal complement subspace $\mathcal{SK}_W$ of the RLD subspace is also dual of $\mathcal{K}_W$ in the following sense. $\mathcal{SK}_W$ is the space of all the tangent vectors which satisfies

$$\mathbf{L}\hat{X} = -(\mathbf{L}\hat{X})^* \tag{22}$$

or, equivalence,

$$(\mathbf{M}\hat{X})W^* + W(\mathbf{M}\hat{X})^* = 0, \tag{23}$$

which yields

$$\pi'_*(\hat{X}) = 0. \tag{24}$$

Metaphorically speaking, (3) and (24) implies that $\mathcal{SK}_W$ looks like $\mathcal{K}_W$ seen from the hidden system.

(22) means that for any member $\hat{X}$ of $\mathcal{SK}_W$, $\pi_*(\hat{X})$ corresponds to a unitary motion of the observed system. The dual of this statement is also valid: for any member $\hat{X}$ of $\mathcal{K}_W$, $\pi'_*(\hat{X})$ corresponds to a unitary motion of the hidden system. This statement reflects the physical fact that the unitary motion of the hidden system do not affect the observed system.

Lemma 1 *The intersection of the $\mathcal{LS}_W$ and $\mathcal{LR}_W$ is given by*

$$\begin{aligned} & \mathcal{LS}_W \cap \mathcal{LR}_W \\ = & \{\hat{X} \mid [\mathbf{L}^S(\pi_*(\hat{X})), \pi(W)] = 0,\ (\mathbf{L}\hat{X})^* = \mathbf{L}\hat{X},\ \} \\ = & \{\hat{X} \mid [\mathbf{L}^S(\pi'_*(\hat{X})), \pi'(W)] = 0,\ (\mathbf{L}\hat{X})^* = \mathbf{L}\hat{X},\ \}. \end{aligned} \tag{25}$$

The intersection of the $\mathcal{K}_W$ and $\mathcal{SK}_W$ is given by

$$\begin{aligned} & \mathcal{K}_W \cap \mathcal{SK}_W \\ = & \{\hat{X} \mid [\mathbf{L}(\hat{X}), \pi(W)] = 0,\ (\mathbf{L}\hat{X})^* = -\mathbf{L}\hat{X},\ \} \\ = & \{\hat{X} \mid [\mathbf{L}'(\hat{X}), \pi'(W)] = 0,\ (\mathbf{L}\hat{X})^* = -\mathbf{L}\hat{X},\ \}, \end{aligned} \tag{26}$$

where $\mathbf{L}'(\hat{X})$ *is defined by*

$$\mathbf{L}'(\hat{X}) \equiv W^{-1}\mathbf{M}(\hat{X})$$

The former statement of the theorem means that for any vector $\hat{X} \in \mathcal{LS}_W \cap \mathcal{LR}_W$, $\pi_*(\hat{X})$ and $\pi'_*(\hat{X})$ correspond to the change of eigenvalues of $\pi(W)$ and $\pi'(W)$, respectively. The latter statement implies that for any vector $\hat{X} \in \mathcal{K}_W \cap \mathcal{SK}_W$, $\pi_*(\hat{X})$ and $\pi'_*(\hat{X})$ correspond to the change of the phase of the eigenvectors of of $\pi(W)$ and $\pi'(W)$, respectively.
<u>Proof</u> (18) yields

$$\pi_*(\hat{X}) = (\mathbf{L}\hat{X})\pi(W) = \pi(W)(\mathbf{L}\hat{X})^*,$$

which, combined with (5), yields

$$[(\mathbf{L}\hat{X}), \pi(W)] = 0$$

Because $\mathbf{L}(\hat{X}) = \mathbf{L}^S(\pi_*(\hat{X})$ holds true for any $\hat{X} \in \mathcal{LS}_W$, we have the first equality in (25). The second equality in (25) and the equalities in (26) are obtained in the same manner. □

9 Canonical distribution

In this section, as an application of the duality of SLD and RLD, we try an estimation theoretical characterization of the canonical model.

One conspicuous feature of the canonical distribution model is that only the eigenvalue of the density matrix is dependent on the parameter, and that the eigenvector is left unchanged even if the parameter changed. Other thermodynamical models, for example, the $T-p$ model

$$\left\{\rho(T,p) \,\middle|\, \rho(T,p) = \sum_{\omega} |\omega\rangle\langle\omega| \exp\left[-\frac{1}{k_B T}\left(E_\omega - pV_\omega - G(T,p)\right)\right]\right\}$$

and the grand canonical model

$$\left\{\rho(T,\mu)\ \middle|\ \rho(T,\mu)=\sum_{\omega}|\omega\rangle\langle\omega|\exp\left[-\frac{1}{k_BT}\left(E_\omega-\mu N_\omega+Y(T,\mu)\right)\right]\right\}$$

also share this feature. Here, H is the Hamiltonian and E_ω the ωth eigenvalue of H, and $|\omega\rangle$ the ωth eigenvector of H. We call the model which has this feature the *classical model.*

Let us require first that the canonical distribution is a pure state in the composite Hilbert space $\mathcal{H}\otimes\mathcal{H}'$, where $\mathcal{H}$ is for the system and $\mathcal{H}'$ for the heat bath (taking trace over the heat bath, we have the canonical distribution). In other words, we assume that the pure state model

$$\mathcal{N}\equiv\{|W(T)\rangle\langle W(T)|\ |\ |W(T)\rangle\in\mathcal{H}\otimes\mathcal{H}'\},\tag{27}$$

the partial trace π over the hidden system $\mathcal{H}'$ reduces to the canonical model $\mathcal{M}$. We denote by $\mathcal{M}'$ the model induced from $\mathcal{N}$ by the partial trace π' over the observed system $\mathcal{H}$.

Second, we assume the following situation: the optimization of the measurement in the estimator of the temperature over any of the following three range

1. all the measurements in $\mathcal{H}$

2. all the measurements in $\mathcal{H}'$

3. all the measurements in $\mathcal{H}\otimes\mathcal{H}'$

achieves exactly the same extent of the efficiency. In usual situation, we can achieve more efficiency in the case of 3 than the other cases, for the range of the optimization is larger. However, as for the macroscopic parameter like the temperature, it is natural to assume that the measurement of the total system do not bring about more information than the measurement of the system. In addition, the measurement of the temperature in the system and the heat bath must yield same amount of information, because they are in the thermal equilibrium. In this situation, we say that the model $\mathcal{M}$ and $\mathcal{N}$ are *maximally entangled.*

Theorem 3 *The model $\mathcal{M}$ and $\mathcal{M}'$ induced by the projection π and π' from the pure state model $\mathcal{N}$ are classical iff they are maximally entangled.*

To prove the theorem, we need the following fact in the pure state estimation theory, which is explained in the later chapters in detail.

Fact The attainable lower bound of the variance of the unbiased estimator of the pure state model (27) is given by $1/\mathrm{tr}A(T)A^*(T)$, where $A(T)$ is the matrix representation of the tangent vector $d/dT|_W \in \mathcal{T}_W(\mathcal{W})$ to the $\mathcal{W}$.

Proof of the theorem Here, we assume $\dim \mathcal{H}'$ is equal to $d = \dim \mathcal{H}$. For the efficiency of the estimation in the case of 1 is equal to that in the case of 3, $d/dT|_W \in \mathcal{T}_W(\mathcal{W})$ need to be a member of the horizontal subspace $\mathcal{LS}_W$, because the length of the horizontal component of $d/dT|_W \in \mathcal{T}_W(\mathcal{W})$ gives the SLD Fisher information of the model $\mathcal{M}$.

Mostly in the same manner, it can be proved that $d/dT|_W \in \mathcal{T}_W(\mathcal{W})$ is the member of $\mathcal{LR}_W$ for the efficiency of the estimation in the case of 2 to be equal to that in the case of 3. Therefore,

$$\left.\frac{d}{dT}\right|_W \in \mathcal{LS}_W \cap \mathcal{LR}_W,$$

which, mixed with lemma 1 leads to the statement of the theorem. □

This theorem, which also applies to the grand canonical model and the $T-p$ model, characterize the entanglement between the heat bath and the system.

Acknowledgment

The author is grateful to Dr. A. Fujiwara, for introducing the quantum estimation theory to the author, and for inspiring discussions.

References

[1] Y. Aharonov and J. Anandan, "Phase change during a cyclic quantum evolution," Phys. Rev. Lett. **58**, 1593–1596 (1987).

[2] S. Amari, *Differential-Geometrical Methods in Statistics*, Lecture Notes in Statistics, Vol. 28 (Springer, Berlin, 1985).

[3] M. V. Berry, "Quantal phase factors accompanying adiabatic changes," Proc. Roy. Soc. London **A392**, 45–57 (1984).

[4] C. W. Helstrom, "Minimum Mean-Square Error Estimation in Quantum Statistics," Phys. Lett., **25A**, 101-102(1967).

[5] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).

[6] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982) (in Russian, 1980).

[7] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry, I, II* (John Wiley, New York, 1963, 1969).

[8] K. Matsumoto, "Uhlmann's parallelity in quantum estimation theory ," METR 97-03 (1997).

[9] H. Nagaoka, "On the Parameter Estimation Problem for Quantum Statistical Models," SITA'89, 577-582 Dec. (1989).

[10] H.Nagaoka, "Differential Geometrical Aspects of Quantum State Estimation and Relative Entropy," METR 94-14 (1994).

[11] J .Samuel and R.Bhandari, "General Setting for Berry's Phase," *Phys.Rev.Lett.***60**, 2239-2342, (1988).

[12] A.Shapere,F.Wilczek, "GEOMETRIC PHASES IN PHYSICS," Advanced Series in Mathematical Physics,vol. 5, World Scientific (1989).

[13] B. Simon, "Holonomy, the quantum adiabatic theorem, and Berry's phase," Phys. Rev. Lett. **51**, 2167–2170 (1983).

[14] A. Uhlmann, "Parallel transport and 'Quantum holonomy' along density operators," Rep. Math. Phys., vol.24, pp.229-240(1986).

[15] A. Uhlmann, “An energy dispersion estimate,” Phys. Lett. A, 161, pp.329–331(1992).

[16] A. Uhlmann, “Density operators as an arena for differential geometry,” Rep. Math. Phys., vol.33, pp.253-263(1993).

[17] H. Yuen and M. Lax, “Multiple-Parameter Quantum Estimation and Measurement of Nnnselfadjoint Observables,” IEEE Trans. on Information Theory, Vol. IT-19, No. 6, 740-750(1973).