http://mmg.tversu.ru ISSN 2311-1275

Mathematical Modelling and Geometry

Volume 4, No 1, p. 19 – 33 (2016)

All the trajectories of an extended averaged Hebbian learning equation on the quantum state space are the e-geodesics

Yoshio Uwano

Department of Mathematics, Kyoto Pharmaceutical University, Misasagi Nakauchi-cho 5, Yamashina-ku, Kyoto, 607-8414, Japan

e-mail: uwano@mb.kyoto-phu.ac.jp

Received 30 January 2016, in final form 14 February. Published 15 February 2016.

Abstract. In this paper, two families of trajectories on the quantum state space (QSS) originating from a synaptic-neuron model and from quantum information geometry meet together. The extended averaged Hebbian learning equation (EAHLE) on the QSS developed by the author and Yuya [1] from a Hebbian synaptic-neuron model is studied from a quantum-information-geometric point of view. It is shown that all the trajectories of the EAHLE are the e-geodesics, the autoparallel curves with respect to the exponential-type parallel transport, on the QSS. As a secondary outcome, an explicit representation of solution of the averaged Hebbian learning equation, the origin of the EAHLE, is derived from that of the e-geodesics on the QSS.

Keywords: dynamical systems, quantum information, geodesic, Hebbian learning

MSC numbers: 37D40, 53C22, 68T05, 81P45

1. Introduction

Quantum computing and quantum information have been well-known to be highly developing fields in which a number of disciplines such as quantum mechanics, mathematics, communication, information, statistics, control, optimization, etc. are crossing over (see [2] for their history, for example). In this paper, a pair of interesting mathematical objects having different origins meet together both of which are described on the quantum state space (QSS), the space of regular density matrices of an arbitrarily fixed degree: One is the extended averaged Hebbian learning equation (EAHLE) on the QSS [1] originating from one of the Hebbian synaptic-neuron learning models and the other is the e-geodesics arising naturally from quantum information geometry.

In order to describe the motive of this paper, a brief history of the series of the author's works [1, 3, 4, 5] is given below. As a departure point of the series of the author's works, Grover's quantum search algorithm [6] is worth touched on, which is well-known to be one of the milestones on the road of quantum computation [2]: For a large number, say $N = 2^n$, of randomly sorted data, the complexity of Grover's algorithm is of $O(\sqrt{N})$, which is lower than the theoretical boundary, $O(N)$, of any non-quantum searches. On Grover's algorithm, Miyake and Wadati [7] made a pioneering geometric study saying that the search sequence is on a geodesic on the $2^{n+1} - 1$ dimensional unit sphere, $S^{2^{n+1}-1}$, of n-qubit states and that the projection of the search sequence on the complex projective space $\mathbf{C}P^{2^n-1}$ is also on a geodesic. Motivated by Miyake and Wadati [7], the author made a geometric study, with Hino and Ishiwatari, on a Grover-type search for an ordered tuple of multi-qubits [3]: As a rigorous analogue to the projection applied to Grover's search sequence, a projection map from the space of ordered tuples to the space of density matrices is constructed. Further, the projection map thus obtained is shown to equip the space of regular density matrices with the SLD-Fisher metric, so that the projection proceeded in [3] is a new geometric realization of the quantum state space (QSS).

A geometric study analogous to [7] about the projection of the Grover-type search on the QSS was however deferred to [5] because of another new interest in the gradient system on the QSS associated with the negative von-Neumann entropy. It is shown in [3] that the gradient system of interest is understood to be a very natural extension of the gradient system associated with the negative Shannon entropy on a classical statistical manifold which is studied by Nakamura [9]. This result encourages the author to seek other noted dynamical systems which are extendable on the QSS. Among the systems displayed in the series of papers [8, 9, 10] by Nakamura on integrable systems, the author and Yuya succeed to extend in [1] the averaged Hebbian learning equation (AHLE) which describes the Hebbian synaptic-neuron learning model proposed by Oja [8, 11]. The dynamical system thus extended on the QSS from the AHLE is the EAHLE dealt with in this paper. A continuous-time limit of Karmarkar's projective scaling algorithm of non-constraint [10, 12] is also shown by the author and Yuya to be extendable on the QSS [4].

After the papers [1, 4] by the author and Yuya, the deferred task for a geometric study of the projection of the Grover-type search sequence for an ordered tuple of multi-qubits is made successfully by the author [5]: The projection of the Grover-type search sequence on the QSS is shown to be on an m-geodesic, an autoparallel curve with respect to the mixture-type parallel transport [13], on the QSS. This result strongly encourages the author to seek other dynamical systems whose trajectories realize geodesics on the QSS.

The aim of this paper is to study the extended averaged Hebbian learning equation (EAHLE) constructed in [1] from a quantum-information-geometric point of view: All the trajectories of the EAHLE are shown to be the e-geodesics, the autoparallel curves with respect to the exponential-type parallel transport [13], on the QSS which are known to play an important role not only in quantum information geometry but also in quantum estimation. In what follows, the organization of this paper is outlined.

At the beginning of the outline, it should be remarked that the pair of sections, Section 2 and Section 3, among five sections in this paper are mostly for reviews of the QSS, the EAHLE and the e-geodesics, which are done according to the author's previous papers [1, 3, 5] and the literature [13] by Hayashi on quantum information. Although the review part seems to occupy rather large part of this paper, it is indispensable because a similarity between looks of the EAHLE and of the tangent vector along the e-geodesics play a key role to reach to the main theorem of this paper.

In Section 2, the QSS is introduced together with the symmetric logarithmic derivative (SLD) and the SLD-Fisher metric on that. The SLD works not only in the definition of the SLD-Fisher metric endowed with the QSS but also in that of the exponential-type parallel transport dealt with in Sec. 3. Section 3 is devoted to reviewing the extended averaged Hebbian learning equation (EAHLE) and the e-geodesics on the QSS. The EAHLE is reviewed in subsection 3.1 together with the way how the EAHLE comes from its origin, the AHLE. In subsection 3.2, the e-geodesics on the QSS are defined to be the autoparallel curves with respect to the exponential-type parallel transport. In Section 4, the main theorem of this paper is proved, which shows that all the trajectories of the EAHLE are the e-geodesics on the QSS. An explicit representation of solution of the AHLE is derived as an outcome of the main theorem. Section 5 is for conclusion.

2. The QSS

In this section, we set up the quantum state space (QSS) as the space of regular density matrices endowed with the SLD-Fisher metric, following [1, 3, 4]. The literature [13] by Hayashi is worth cited to have a general framework on quantum information geometry including the QSS, in which the QSS is referred to as 'the space of quantum states'.

Let Q_n be the set of $n \times n$ regular density matrices, namely, the set of $n \times n$ positive-definite Hermitean matrices with unit trace. On denoting by $M(n)$ the set of all the $n \times n$ complex matrices, Q_n is defined to be the set

$$Q_n = \left\{ \rho \in M(n) \,\middle|\, \rho : \text{positive definite},\ \rho^\dagger = \rho,\ \mathrm{Tr}\,\rho = 1 \right\}, \tag{1}$$

where † stands for the Hermitean conjugate operation and Tr for the trace of matrices. The tangent space, denoted by $T_\rho Q_n$, of Q_n at $\rho \in Q_n$ then takes the form

$$T_\rho Q_n = \left\{ X \in M(n) \,\middle|\, X^\dagger = X,\ \mathrm{Tr}\,X = 0 \right\}, \tag{2}$$

which is equipped with the $\mathbf{R}$-vector space structure.

As a natural quantum-information-geometric structure of Q_n, the SLD-Fisher metric is endowed with Q_n in what follows. For the endowment, we need to introduce the symmetric logarithmic derivative (SLD) on tangent vectors. The SLD, denoted by $L_\rho(X)$, on $X \in T_\rho Q_n$ is the $n \times n$ matrix determined uniquely by the equation

$$X = \frac{1}{2}\{\rho L_\rho(X) + L_\rho(X)\,\rho\} \quad (X \in T_\rho Q_n). \tag{3}$$

It follows from (3) that the SLD satisfies

$$(L_\rho(X))^\dagger = L_\rho(X) \quad (X \in T_\rho Q_n). \tag{4}$$

The matrix-element display of the SLD given in [1, 3, 4] is of great help also in this paper, which is utilized to prove our main theorem in Sec.4. Let $\rho \in Q_n$ be written in the form

$$\rho = h\,\mathrm{diag}\,(\theta_1, \theta_2 \cdots, \theta_n)\; h^\dagger \quad (h \in U(n)), \tag{5}$$

where $\mathrm{diag}\,(\theta_1, \theta_2 \cdots, \theta_n)$ denotes the $n \times n$ diagonal matrix whose j-th diagonal entry is θ_j $(j = 1, 2, \cdots, n)$ and $U(n)$ stands for the group of $n \times n$ unitary matrices. The symbol 'diag' indicates the diagonal matrices henceforce. On expressing $X \in T_\rho Q_n$ as

$$X = h\widetilde{X}h^\dagger \tag{6}$$

with $h \in U(n)$ of (5), the (j,k)-entry of $h^\dagger L_\rho(X) h$ $(X \in T_\rho Q_n)$ is calculated to be

$$\left(h^\dagger L_\rho(X) h\right)_{jk} = \left(\frac{2}{\theta_j + \theta_k}\right)\widetilde{X}_{jk} \quad (j, k = 1, 2, \cdots, n). \tag{7}$$

Equations (3)-(7) are put together to show the following lemma on the SLD.

Lemma 1. *The symmetric logarithmic derivative (SLD), L_ρ, is a one-to-one and onto $\mathbf{R}$-linear map from $T_\rho Q_n$ to*

$$L_\rho(T_\rho Q_n) = \left\{ \Xi \in M(n) \,\middle|\, \Xi^\dagger = \Xi,\ \mathrm{Tr}\,(\rho\Xi + \Xi\rho) = 0 \right\} \quad (\rho \in Q_n). \tag{8}$$

Under (5) and

$$\Xi = h\widetilde{\Xi}h^\dagger \in L_\rho(T_\rho Q_n), \tag{9}$$

the inverse, denoted by L_ρ^{-1}, *of the SLD is given to satisfy*

$$\left(h^\dagger L_\rho^{-1}(\Xi)h\right)_{jk} = \left(\frac{\theta_j + \theta_k}{2}\right)\widetilde{\Xi}_{jk} \quad (j,k = 1,2,\cdots,n). \tag{10}$$

In terms of the SLD, the SLD-Fisher metric $\langle\cdot,\cdot\rangle$ is defined by

$$\langle X,\, X'\rangle_\rho = \mathrm{Tr}\,\left(X^\dagger L_\rho(X')\right) \quad (X, X' \in T_\rho Q_n) \tag{11}$$

(see Hayashi [13]). On using (3) and (4), the SLD-Fisher metric is brought into the form

$$\langle X,\, X'\rangle_\rho = \frac{1}{2}\,\mathrm{Tr}\,\left(\rho\left(L_\rho(X)L_\rho(X') + L_\rho(X')L_\rho(X)\right)\right) \quad (X, X' \in T_\rho Q_n) \tag{12}$$

[1, 3, 4, 5]. Furthermore, with the matrix-element displays, (5)-(7), and

$$X' = h\widetilde{X}'h^\dagger, \tag{13}$$

the SLD-Fisher metric is expressed to be

$$\langle X, X'\rangle_\rho = \sum_{j,k=1}^{n}\left(\frac{2}{\theta_j + \theta_k}\right)\overline{\widetilde{X}}_{jk}\widetilde{X}'_{jk} \tag{14}$$

[1, 3, 4, 5]. Equation (14) works effectively to derive a useful formula to the gradient equation on the QSS [1, 3, 4]. The Riemannian manifold Q_n endowed with the SLD-Fisher metric $\langle\cdot,\cdot\rangle$ is what we are referring to as the quantum state space.

3. The EAHLE and the e-geodesics

In this section, the extended averaged Hebbian learning equation (EAHLE) and the e-geodesics are introduced according to [4] for the EAHLE and to [13] for the e-geodesics.

3.1 The EAHLE

The extended averaged Hebbian learning equation (EAHLE) is organized by the present author and Yuya [4] who are inspired by Nakamura's paper [8] on the averaged Hebbian learning equation (AHLE). According to [4], the EAHLE is the first order differential equation

$$\frac{d\rho}{dt} = \rho C + C\rho - 2\,\mathrm{Tr}\,(C\rho)\,\rho \tag{15}$$

on the QSS. The C on the rhs of (15) is the real diagonal matrix

$$C = \mathrm{diag}\,(c_1, c_2, \cdots, c_n) \tag{16}$$

of degree n, whose diagonal entries stand for the eigenvalues of the autocorrelation matrix of the stationary stochastic process governing the Hebbian learning process [8, 11, 14].

The reason for referring to Eq. (15) as the 'extended' averaged Hebbian learning equation is given in what follows. Let $w = (w_j)_{j=1,2,\cdots,n}$ be the variables on t expressing the coupling strengths of neurons obtained through an appropriate change of the independent variable to have t. The first-order differential equation

$$\frac{dw}{dt} = Cw - \left(w^T C w\right)\, w \tag{17}$$

on the $n-1$ dimensional sphere

$$S^{n-1} = \left\{ w = (w_1, w_2, \cdots, w_n)^T \in \mathbf{R}^n \;\middle|\; w^T w = 1 \right\} \tag{18}$$

with unit radius describes Oja's rule [11] on the Hebbian learning process of synaptic neurons [14], where C is the diagonal matrix given by (16). In (17), the C is understood again to be the diagonalization of the autocorrelation matrix of the governing stationary stochastic process of neurons. The differential equation (17) is what we are referring to as the averaged Hebbian learning equation (AHLE).

If we restrict Eq. (17) on each of the open subsets,

$$\begin{aligned} S^{n-1}_\sigma = \left\{ w \in S^{n-1} \;\middle|\; \sigma_j w_j > 0,\ j = 1, 2, \cdots, n \right\} \\ (\sigma = (\sigma_j),\ \sigma_j = \pm 1,\ j = 1, 2, \cdots, n)\,, \end{aligned} \tag{19}$$

of S^{n-1}, it is brought into the first-order differential equation

$$\frac{d\theta_j}{dt} = 2c_j\theta_j - 2\left(\sum_{k=1}^{n} c_j\theta_j\right)\theta_j \quad (j = 1, 2, \cdots, n), \tag{20}$$

on the submanifold

$$D_n = \{\Theta \in Q_n \mid \Theta = \mathrm{diag}\,(\theta_1, \theta_2, \cdots, \theta_n)\,\} \tag{21}$$

of Q_n through the map

$$\begin{aligned} p_{n,\sigma}(w) = \mathrm{diag}\left(w_1^2, w_2^2, \cdots, w_n^2\right) \\ \left(w \in S^{n-1}_\sigma,\ \sigma = (\sigma_j),\ \sigma_j = \pm 1,\ j = 1, 2, \cdots, n\right) \end{aligned} \tag{22}$$

from S^{n-1}_σ to D_n [1]. The θ_j's in (20) and (21) are subject to the constraints

$$\theta_j > 0 \; (j = 1, 2, \cdots n) \quad \text{and} \quad \sum_{j=1}^{n} \theta_j = 1. \tag{23}$$

We note here that Eq. (20) is the same form as the Toda lattice written in Moser's form [8, 15, 16].

Since every mapping $p_{n,\sigma}$ defined by (22) is a diffeomorphism with the inverse

$$p_{n,\sigma}^{-1}(\Theta) = \left(\sigma_1\sqrt{\theta_1}, \sigma_2\sqrt{\theta_2}, \cdots, \sigma_n\sqrt{\theta_n}\right)$$
$$(\Theta \in D_n,\ \sigma = (\sigma_j),\ \sigma_j = \pm 1,\ j = 1, 2, \cdots, n)\,, \tag{24}$$

we can understand that the differential equation (20) on D_n is a 'copy' of the AHLE restricted on each S_σ^n and vice versa [1].

We are at the final stage to account for the naming of the EAHLE. To complete the account, we show that the differential equation (20) on D_n is the restriction of the EAHLE, (15), on D_n. In fact, the substitution of $\Theta \in D_n$ for ρ in (15) yields the 'copy', (20), of the AHLE. In a summary, we have the following lemma [1, 8].

Lemma 2. *The restriction of the averaged Hebbian learning equation (AHLE) on each S_σ^{n-1} is equivalent, up to the diffeomorphisms given by (22), to Eq. (20) which describes not only the restriction of the extended averaged Hebbian learning equation (EAHLE) on D_n but also the Toda lattice in Moser's form [8, 15, 16].*

3.2 The e-geodesics on the QSS

To those who are not familiar with differential geometry, a geodesic connecting a given pair of points would be thought of as the shortest-distance path between the given points. For example, in the Euclidean plane, a typical model space for school-geometry, we are taught that the straight-line segment connecting a given pair of points is the geodesic between them. In differential geometry, however, the notion of length or distance, is unnecessary in defining geodesics: What is needed in the definition of geodesics is the idea of parallel transports, namely, the idea for comparing tangent vectors at a certain point with those at another point. Once a parallel transport is fixed, the geodesics are defined to be the autoparallel curves with respect to that parallel transport. For an intuitive description of geodesics and parallel transports, the literature [17] by Nakahara is worth cited.

Let us start with the definition of the exponential-type (e-) parallel transport. According to Hayashi [13], the e-parallel transport from $T_{\rho_1}Q_n$ to $T_{\rho_2}Q_n$ is the $\mathbf{R}$-linear map $\tau_{\rho_1,\rho_2} : T_{\rho_1}Q_n \to T_{\rho_2}Q_n$ subject to

$$L_{\rho_2}(\tau_{\rho_1,\rho_2}(X)) = L_{\rho_1}(X) - \mathrm{Tr}\,(\rho_2\, L_{\rho_1}(X)) \quad (X \in T_{\rho_1}Q_n), \tag{25}$$

where L_{ρ_1} and L_{ρ_2} denote the SLD defined by (3) with $\rho = \rho_1$ and $\rho = \rho_2$, respectively. Combining the defining equation (3) for the SLD with Eq. (25), we can obtain a more direct form,

$$\tau_{\rho_1,\rho_2}(X) = \frac{1}{2}\left\{\rho_2 L_{\rho_1}(X) + L_{\rho_1}(X)\rho_2\right\} - \mathrm{Tr}\,(\rho_2\, L_{\rho_1}(X))\,\rho_2 \quad (X \in T_{\rho_1}Q_n), \tag{26}$$

of the e-parallel transport. The e-parallel transport satisfies, of course, the postulate of parallel transports (see Guggenheimer [18], for example), but we do not get it into detail here.

Definition 1. *Tangent vectors $X_1 \in T_\rho Q_n$ and $X_2 \in T_{\rho_2} Q_n$ are e-parallel if they are parallel with respect to the e-parallel transport; namely, X_1 and X_2 are e-parallel if they satisfy*

$$X_2 = \tau_{\rho_1,\rho_2}(X_1), \tag{27}$$

where τ_{ρ_1,ρ_2} is the e-parallel transport given by (26).

Once we fix a parallel transport, we can consider the geodesics as the autoparallel curves to that parallel transport [17, 18]. In view of Definition 1, we can define the e-geodesics as follows [13].

Definition 2. *A smooth curve $\rho(t)$ $(0 \le t \le {}^\exists T)$ on the QSS is an e-geodesic if it satisfies*

$$\begin{aligned}\frac{d\rho}{dt}(t) &= \tau_{\rho(0),\,\rho(t)}\left(\frac{d\rho}{dt}(0)\right)\\ &= \frac{1}{2}\left\{\rho(t)L_{\rho(0)}\left(\frac{d\rho}{dt}(0)\right) + L_{\rho(0)}\left(\frac{d\rho}{dt}(0)\right)\rho(t)\right\}\\ &\qquad - \mathrm{Tr}\left(L_{\rho(0)}\left(\frac{d\rho}{dt}(0))\right)\rho(t)\right)\rho(t) \qquad (0 \le t \le T),\end{aligned} \tag{28}$$

where $\tau_{\rho(0),\,\rho(t)}$ is the e-parallel transport from $T_{\rho(0)}Q_n$ to $T_{\rho(t)}Q_n$ given by (26) with $\rho(0)$ and $\rho(t)$ in place of ρ_1 and ρ_2, respectively.

Equation (28) for the autoparallelism with respect to the e-parallel transport (26) must not be understood to be a first-order differential equation on the QSS because of the appearance of the initial tangent vector $(d\rho/dt)(0)$ in the rhs of (28) that never takes place in first-order differential equations. Hence the expression (28) of the e-geodesics does not contradict the second-order-differential-equation form taught in theory of geodesics. According to Hayashi [13], the e-geodesic admits the explicit representation below.

Lemma 3. *The e-geodesic $\rho^e\left(t; \rho^{(0)}, X^{(0)}\right)$ with the initial conditions,*

$$\rho^e\left(0; \rho^{(0)}, X^{(0)}\right) = \rho^{(0)} \in Q_n \tag{29}$$

and

$$\frac{d\rho^e}{dt}\left(0; \rho^{(0)}, X^{(0)}\right) = X^{(0)} \in T_{\rho^{(0)}}Q_n, \tag{30}$$

takes the form

$$\begin{aligned}\rho^e\left(t; \rho^{(0)}, X^{(0)}\right) = &\left\{\mathrm{Tr}\left(e^{\frac{t}{2}L_{\rho^{(0)}}(X^{(0)})}\rho^{(0)}\, e^{\frac{t}{2}L_{\rho^{(0)}}(X^{(0)})}\right)\right\}^{-1}\\ &\times e^{\frac{t}{2}L_{\rho^{(0)}}(X^{(0)})}\rho^{(0)}\, e^{\frac{t}{2}L_{\rho^{(0)}}(X^{(0)})}\end{aligned} \tag{31}$$

for $0 \le t < \infty$, where $L_{\rho^{(0)}}\left(X^{(0)}\right)$ is the SLD, defined by (3), on $X^{(0)} \in T_{\rho^{(0)}}Q_n$.

A direct differentiation of (31) by t clearly shows that $\rho^e(t;\rho^{(0)},X^{(0)})$ satisfies Eq. (28) of the autoparallelism. Namely, we have

$$\begin{aligned}&\frac{d\rho^e}{dt}\left(t;\rho^{(0)},X^{(0)}\right)\\ =&\frac{1}{2}\left\{\rho^e\left(t;\rho^{(0)},X^{(0)}\right)L_{\rho^{(0)}}\left(X^{(0)}\right)+L_{\rho^{(0)}}\left(X^{(0)}\right)\rho^e\left(t;\rho^{(0)},X^{(0)}\right)\right\}\\ &-\mathrm{Tr}\left(L_{\rho^{(0)}}\left(X^{(0)}\right)\rho^e\left(t;\rho^{(0)},X^{(0)}\right)\right)\rho^e\left(t;\rho^{(0)},X^{(0)}\right).\end{aligned}\tag{32}$$

We see that Eq. (32) for any fixed e-geodesic looks quite similar to the EAHLE (15) and therefore we may expect that every trajectory of the EAHLE can be realized as an e-geodesic.

4. The EAHLE trajectories as the e-geodesics

Now that we have found the similarity of the EAHLE (15) and Eq. (32) for the e-geodesics, we show that the trajectories of the EAHLE are the e-geodesics below. Further, as an outcome of Theorem 1, an explicit representation of solution of the AHLE is derived from the representation (31) of the e-geodesics.

On comparing very naively Eq. (15) with Eq. (32), one might come to choose $L_{\rho^{(0)}}(X^{(0)})$ in (32) to be equal to C in (15). However, this choice fails because the diagonal matrix C never belongs to $L_{\rho^{(0)}}(T_{\rho^{(0)}}Q_n)$ (see (8) with $\rho^{(0)}$ in place of ρ). By choosing $L_{\rho^{(0)}}(X^{(0)})$ in (32) suitably, we have the following theorem to characterize all the trajectories of the EAHLE as the e-geodesics.

Theorem 1 (Main Theorem). *For any fixed $\rho^{(0)}\in Q_n$, let $\rho^h(t;\rho^{(0)})$ denote the trajectory of the EAHLE subject to the initial condition*

$$\rho^h(0;\rho^{(0)})=\rho^{(0)},\tag{33}$$

and let $\rho^e(t;\rho^{(0)},X^{(C)})$ denote the e-geodesic $\rho^e(t;\rho^{(0)},X^{(C)})$ subject to the initial conditions

$$\rho^e(0;\rho^{(0)},X^{(C)})=\rho^{(0)}\tag{34}$$

and

$$\frac{d\rho^e}{dt}(0;\rho^{(0)},X^{(C)})=X^{(C)}\tag{35}$$

with

$$X^{(C)}=\rho^{(0)}C+C\rho^{(0)}-2\mathrm{Tr}\,(C\rho^{(0)})\rho^{(0)}.\tag{36}$$

Then, for $t\geq 0$, the trajectory $\rho^h(t;\rho^{(0)})$ of the EAHLE coincides with the e-geodesic $\rho^e(t;\rho^{(0)},X^{(C)})$.

Proof From (33) and (34), we easily confirm that $\rho^h(t;\rho^{(0)})$ and $\rho^e(t;\rho^{(0)},X^{(C)})$ share $\rho^{(0)}$ as the initial point. As a necessary condition for $\rho^e(t;\rho^{(0)},X^{(C)}) = \rho^h(t;\rho^{(0)})$, we pose the coincidence

$$\frac{d\rho^e}{dt}(0;\rho^{(0)},X^{(C)}) = \frac{d\rho^h}{dt}(0;\rho^{(0)}) \tag{37}$$

of the initial tangent vectors, which is equivalent to the pair of equations, (35) and (36). Hence, what we have to show here is that the pair, (35) and (36), is also a sufficient condition for $\rho^e(t;\rho^{(0)},X^{(C)}) = \rho^h(t;\rho^{(0)})$. Through the proof, we often apply the abbreviation ρ^e to $\rho^e(t:\rho^{(0)},X^{(C)})$. The core part of the proof is given in what follows in a straightforward calculation form with the notation

$$h^\dagger \rho^{(0)} h = \Theta^{(0)} = \mathrm{diag}\,(\theta_1^{(0)},\theta_2^{(0)},\cdots,\theta_n^{(0)}) \quad (h \in U(n)) \tag{38}$$

and

$$\widetilde{X}^{(C)} = h^\dagger X^{(C)} h, \quad \widetilde{L}^{(0)} = h^\dagger L_{\rho^{(0)}}(X^{(C)}) h, \quad \widetilde{C} = h^\dagger C h, \quad R = h^\dagger \rho^e h. \tag{39}$$

We note here that $\widetilde{C}$ and R are non-diagonal in general and that R is of unit trace. With the notation given above, the matrix-element display of $\widetilde{X}^{(C)}$ takes the form

$$\begin{aligned}\widetilde{X}^{(C)}_{jk} &= \left(h^\dagger\{\rho^{(0)}C + C\rho^{(0)} - 2\,\mathrm{Tr}\,(C\rho^{(0)})\rho^{(0)}\}h\right)_{jk} \\ &= \left(\Theta^{(0)}\widetilde{C} + \widetilde{C}\Theta^{(0)} - 2\,\mathrm{Tr}\,(\widetilde{C}\Theta^{(0)})\Theta^{(0)}\right)_{jk} \\ &= \theta_j^{(0)}\widetilde{C}_{jk} + \widetilde{C}_{jk}\theta_k^{(0)} - 2\,\mathrm{Tr}\,(\widetilde{C}\Theta^{(0)})\delta_{jk}\theta_k^{(0)} \quad (j,k=1,2,\cdots,n),\end{aligned} \tag{40}$$

where the symbol δ_{jk} indicates Kronecker's delta $(j,k=1,2,\cdots,n)$. Equation (40) is combined with (7) to yield the matrix-element display

$$\widetilde{L}^{(0)}_{jk} = 2\widetilde{C}_{jk} - \left(\frac{4}{\theta_j^{(0)}+\theta_k^{(0)}}\right)\,\mathrm{Tr}\,\left(\widetilde{C}\Theta^{(0)}\right)\delta_{jk}\theta_k^{(0)} \quad (j,k=1,2,\cdots,n) \tag{41}$$

for $\widetilde{L}^{(0)}$.

We are now in a position to calculate the rhs of (32) with $X^{(0)} = X^{(C)}$ and (36). Putting Eqs. (38), (39) and (41) together with the abbreviation ρ^e for

$\rho^e(t;\rho^{(0)},X^{(C)})$, we have

$$
\begin{aligned}
&\left(h^\dagger\left\{\frac{1}{2}\left(\rho^e L_{\rho^{(0)}}\left(X^{(C)}\right)+L_{\rho^{(0)}}\left(X^{(C)}\right)\rho^e\right)-\mathrm{Tr}\,\left(L_{\rho^{(0)}}(X^{(C)})\rho^e\right)\rho^e\right\}h\right)_{jk}\\
=&\left(\frac{1}{2}\left(R\widetilde{L}^{(0)}+\widetilde{L}^{(0)}R\right)-\mathrm{Tr}\,\left(\widetilde{L}^{(0)}R\right)R\right)_{jk}\\
=&\frac{1}{2}\sum_{m=1}^{n}\left\{\left(R_{jm}\widetilde{L}^{(0)}_{mk}+\widetilde{L}^{(0)}_{jm}R_{mk}\right)-2\mathrm{Tr}\,(\widetilde{L}^{(0)}R)R_{jk}\right\}\\
=&\sum_{m=1}^{n}R_{jm}\left\{\widetilde{C}_{mk}-\left(\frac{2}{\theta^{(0)}_m+\theta^{(0)}_k}\right)\mathrm{Tr}\,\left(\widetilde{C}\Theta^{(0)}\right)\delta_{mk}\theta^{(0)}_k\right\}\\
&\quad-\sum_{m=1}^{n}\left\{\widetilde{C}_{jm}-\left(\frac{2}{\theta^{(0)}_j+\theta^{(0)}_m}\right)\mathrm{Tr}\,\left(\widetilde{C}\Theta^{(0)}\right)\delta_{jm}\theta^{(0)}_m\right\}R_{mk}\\
&\qquad-2\left[\sum_{m,l=1}^{n}\left\{\widetilde{C}_{lm}-\left(\frac{2}{\theta^{(0)}_l+\theta^{(0)}_m}\right)\mathrm{Tr}\,\left(\widetilde{C}\Theta^{(0)}\right)\delta_{lm}\theta^{(0)}_m\right\}R_{ml}\right]R_{jk}\\
=&\sum_{m=1}^{n}\left(R_{jm}\widetilde{C}_{mk}+\widetilde{C}_{jm}R_{mk}\right)-2\mathrm{Tr}\,\left(\widetilde{C}\Theta^{(0)}\right)R_{jk}\\
&\quad-2\left(\sum_{m,l=1}^{n}\widetilde{C}_{lm}R_{ml}\right)R_{jk}+2\mathrm{Tr}\,\left(\widetilde{C}\Theta^{(0)}\right)\left(\sum_{m=1}^{n}R_{mm}\right)R_{jk}\\
=&\left(R\widetilde{C}+\widetilde{C}R\right)_{jk}-2\mathrm{Tr}\,\left(\widetilde{C}R\right)R_{jk}\\
=&\left(h^\dagger\left\{\rho^e C+\rho^e C-2\mathrm{Tr}\,(C\rho^e)\rho^e\right\}h\right)_{jk}.
\end{aligned}
\tag{42}
$$

Equation (42) is put together with Eq. (32) to show that $\rho^e(t:\rho^{(0)},X^{(C)})$ satisfies the equation

$$
\begin{aligned}
\frac{d\rho^e}{dt}(t:\rho^{(0)},X^{(C)})=&\rho^e\left(t:\rho^{(0)},X^{(C)}\right)C+C\rho^e\left(t:\rho^{(0)},X^{(C)}\right)\\
&-2\mathrm{Tr}\,(C\rho^e(t:\rho^{(0)},X^{(C)}))\rho^e(t:\rho^{(0)},X^{(C)}).
\end{aligned}
\tag{43}
$$

which turns out to be the same as Eq. (15) with $\rho^e(t;\rho^{(0)},X^{(C)})$ in place of ρ. Put in another way, the e-geodesic $\rho^e(t;\rho^{(0)},X^{(C)})$ satisfies the EAHLE (15) with the initial condition $\rho(0)=\rho^{(0)}$ and accordingly $\rho^e(t;\rho^{(0)},X^{(C)})$ coincides with $\rho^h(t;\rho^{(0)})$. This completes the proof.

Combining Lemma 2, Lemma 3 and Theorem 1 together, we can give an explicit representation of solution of the averaged Hebbian learning equation (AHLE) from the representation, (31), of the e-geodesics. Although it is pointed out in Nakamura [8] that the representation of solution of the AHLE is available from that of the Toda lattice in Moser's form [15, 16] (see also Lemma 2), we are to present the same one as in [8] because our derivation process below is new.

On recalling Lemma 2, the representation of solution of the AHLE can be given by calculating explicitly the Eq. (31) with the initial conditions,

$$\rho^e(0) = \rho^{(0)} = \Theta^{(0)} = \mathrm{diag}\,\left(\theta_1^{(0)}, \theta_2^{(0)}, \cdots, \theta_n^{(0)}\right) \in D_n \subset Q_n \tag{44}$$

and

$$\frac{d\rho^e}{dt}(0) = X^{(0)} = \Xi^{(C)} = \mathrm{diag}\,\left(\xi_1^{(0)}, \xi_2^{(0)}, \cdots, \xi_n^{(0)}\right) \in T_{\Theta^{(0)}}D_n \subset T_{\Theta^{(0)}}Q_n \tag{45}$$

with

$$\xi_j^{(0)} = 2c_j\theta_j^{(0)} - 2\mathrm{Tr}\,\left(C\Theta^{(0)}\right)\theta_j^{(0)} \qquad (j = 1, 2, \cdots, n), \tag{46}$$

where C is the diagonal matrix governing both the EAHLE (15) and the AHLE (17). Under (44)-(46), the SLD of $\Xi^{(C)}$ and its exponential are calculated to be

$$L_{\Theta^{(0)}}\left(\Xi^{(C)}\right) = \mathrm{diag}\,\left(\frac{\xi_1^{(0)}}{\theta_1^{(0)}}, \frac{\xi_2^{(0)}}{\theta_2^{(0)}}, \cdots, \frac{\xi_n^{(0)}}{\theta_n^{(0)}}\right) = 2C - 2\mathrm{Tr}\,\left(C\Theta^{(0)}\right)I \tag{47}$$

and

$$e^{\frac{t}{2}L_{\Theta^{(0)}}\left(\Xi^{(C)}\right)} = e^{-t\,\mathrm{Tr}\,\left(C\Theta^{(0)}\right)}\,\mathrm{diag}\,\left(e^{tc_1}, e^{tc_2}, \cdots, e^{tc_n}\right). \tag{48}$$

Hence it follows from Lemma 3 and Theorem 1 that the trajectory on D_n, denoted by $\Theta^h(t)$, of the EAHLE with the initial condition (44) takes the form

$$\Theta^h(t) = \mathrm{diag}\,\left(\theta_1^h(t), \theta_2^h(t), \cdots, \theta_n^h(t)\right) \tag{49}$$

with

$$\theta_j^h(t) = \left(\sum_{k=1}^n e^{2tc_k}\theta_k^{(0)}\right)^{-1} e^{2tc_1}\theta_j^{(0)} \quad (j = 1, 2, \cdots, n). \tag{50}$$

We note here that Eqs. (49) and (50) reproduce the solution of the Toda lattice in Moser's form [15, 16] in view of Lemma 2. Applying the map $p_{n,\sigma}^{-1}$ defined by (24) to $\Theta^h(t)$, we have

$$\begin{aligned} w_\sigma^h(t) &= p_{n,\sigma}^{-1}(\Theta^h(t)) \\ &= \left(\sum_{k=1}^n e^{2tc_k}\theta_k^{(0)}\right)^{-1/2} \left(e^{tc_1}\sigma_1\sqrt{\theta_1^{(0)}}, e^{tc_2}\sigma_2\sqrt{\theta_2^{(0)}}, \cdots, e^{tc_n}\sigma_n\sqrt{\theta_n^{(0)}}\right)^T \\ &\qquad (\sigma = (\sigma_j),\ \sigma_j = \pm 1,\ j = 1, 2, \cdots, n) \end{aligned} \tag{51}$$

which realizes the solutions of the AHLE on the open-dense subset $\cup_\sigma S_\sigma^{n-1}$ of S^{n-1} (see (19) for S_σ^{n-1}). Thus we have the following corollary to Theorem 1.

Corollary 1. *The solution of the averaged Hebbian learning equation (17) subject to the initial condition $w(0) = (w_1^{(0)}, w_2^{(0)}, \cdots, w_n^{(0)})^T \in S^{n-1}$ is given by*

$$w^h(t) = \left(\sum_{k=1}^n e^{2tc_k}\left(w_k^{(0)}\right)^2\right)^{-1/2} \left(e^{tc_1}w_1^{(0)}, e^{tc_2}w_j^{(0)}, \cdots, e^{tc_n}w_n^{(0)}\right)^T. \tag{52}$$

5. Conclusions

We show in Theorem 1 that all the trajectories of the extended averaged Hebbian learning equation (EAHLE) on the QSS are the e-geodesics on the QSS. As a direct application of Theorem 1, the explicit representation, (52), of solution of the averaged Hebbian learning equation (AHLE), the departure equation of the EAHLE, is derived from the representation, (31), of the e-geodesics. Although the expression (52) is known already to be available from the Toda lattice in Moser's form due to the equivalence between the AHLE and the Moser's form [8], our derivation in Sec.4 is worth given because it is made from a novel point of view, a quantum-information-geometric point of view.

We would like to offer a remark on Theorem 1 from a geometric-mechanics point of view: In view of the gradient-system structure of the EAHLE revealed in Uwano and Yuya [1], Theorem 1 is understood to provide a gradient system whose trajectories are the e-geodesics. Further, since the e-geodesics are known to play an important role in quantum estimation [13], the EAHLE is expected to be a new candidate of gradient systems dealt with in Braunstein [19] for quantum estimation. Another remark is offered from an integrable-systems point of view: The EAHLE would be looked on as an extended Toda lattice in Moser's form: This view is supported from the coincidence given in Lemma 2 between the EAHLE restricted on D_n and the Moser's form.

On closing this paper, the author would like to make the following conjecture on the e-geodesics which do not satisfy the initial conditions, (35) and (36), attached to the initial tangent vector.

Conjecture *All the e-geodesics on the QSS are realized as the EAHLE trajectories up to the adjoint $SU(n)$ actions on the QSS and the affine transformations of time.*

The conjecture will be investigated soon together with the dynamics on the QSS described by the EAHLE.

Acknowledgements. The author thanks Professor Yoshimasa Nakamura at Kyoto University for his valuable remark on the paper [8] and his suggestion to include our derivation of the AHLE solution (Corollary 1) in the present paper.

References

[1] Uwano Y. and Yuya H. *A Hebb-type learning equation on the quantum information space – A clue to a fast principal component analyzer.* Far East Journal of Applied Mathematics 2010, **47**, pp. 149-167

[2] Nielsen M. A. and Chuang I. L. *Quantum Computation and Quantum Information.* 2000, Cambridge: Cambridge U.P.

[3] Uwano Y., Hino H. and Ishiwatari Y. *Certain integrable system on a space associated with a quantum search algorithm.* Physics of Atomic Nuclei 2007, **70**, pp. 784-791

[4] Uwano Y. and Yuya H. *A gradient system on the quantum information space realizing the Karmarkar flow for linear programming – A clue to effective algorithms –.* New Trends in Quantum Information (Special Issue of Electronic Journal of Theoretical Physics) 2010, Rome: Aracne Editorice, pp. 257-276

[5] Uwano Y. *Geometry of Dynamics of a Quantum Search Algorithm for an Ordered Tuple of Multi-Qubits.* Search Algorithms for Engineering Optimization 2013, Riejka: InTech Press, pp. 261-284

[6] Grover L. *A fast quantum mechanical algorithm for database search.* 1996, In: Proc. the 28th Annual ACM Symposium on the Theory of Computing, pp. 212-219, New York: ACM

[7] Miyake A. and Wadati M. *Geometric strategy for the optimal quantum search.* Phys. Rev. A 2001, **64**, 042317, 9pp.

[8] Nakamura Y. *Neurodyanamics and nonlinear integrable systems of Lax type.* Japan J. Indust. Appl. Math. 1994, **11**, pp. 11-20

[9] Nakamura Y. *Completely integrable dynamical systems on the manifolds of Gaussian and multinomial distributions.* Japan J. Indust. Appl. Math. 1993, **10**, pp. 179-189

[10] Nakamura Y. *Lax pair and fixed point analysis of Karmarkar's projective scaling trajectory for linear programming.* Japan J. Indust. Appl. Math. 1994, **11**, pp. 1-9

[11] Oja E. *A simplified neuron model as a principal component analyzer.* J. Math. Biology 1982, **15**, pp. 267-273

[12] Karmarkar N. *Riemannian geometry underlying interior-point methods for linear programming.* Mathematical Developments Arising from Linear Programming (Contemp. Math. **114**) 1990, Providence: AMS, pp. 51-75

[13] Hayashi M. *Quantum Information.* 2006, Berlin: Springer-Verlag

[14] Hebb D. O. *The Organization of Behavior.* 1949, New York: Wiley

[15] Moser J. *Finitely many mass points on the line under the influence of an exponential potential – an integrable system.* Dynamical Systems, Theory and Applications (LNP **38**) 1975, Heidelberg: Springer-Verlag, pp. 467-497

[16] Moser J. and Zehnder E.J. *Notes on Dynamical Systems* 2005, Providence: AMS

[17] Nakahara M. *Geometry, Topology and Physics.* 1990, Bristol: IOP Publishing

[18] Guggenheimer H. *Differential Geometry.* 1977, New York: Dover

[19] Braunstein S.L. *Geometry of quantum inference.* Phys. Lett. A 1996, **219**, pp. 169-174