Information and contact geometric description of expectation variables exactly derived from master equations

Shin-itiro Goto and Hideitsu Hino

The Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan

E-mail: sgoto@ism.ac.jp

Received 2 April 2019, revised 20 July 2019
Accepted for publication 9 September 2019
Published 20 December 2019

Abstract

In this paper a class of dynamical systems describing expectation variables exactly derived from continuous-time master equations is introduced and studied from the viewpoint of differential geometry, where such master equations consist of a set of appropriately chosen Markov kernels. To geometrize such dynamical systems for expectation variables, information geometry is used for expressing equilibrium states, and contact geometry is used for nonequilibrium states. Here time-developments of the expectation variables are identified with contact Hamiltonian vector fields on a contact manifold. Also, it is shown that the convergence rate of this dynamical system is exponential. Duality emphasized in information geometry is also addressed throughout.

Keywords: master equations, information geometry, contact geometry, statistical physics

1. Introduction

Information geometry is a geometrization of mathematical statistics [1], and its structure and applications have been investigated. This geometry offers tools to study statistical quantities defined on statistical manifolds, where statistical manifolds are identified with parameter spaces for parametric distribution functions. Examples of applications of information geometry include statistical interference, quantum information, and thermodynamics [2–6]. From these examples one sees that the application of information geometry to sciences and engineering enables one to visualize theories and to utilize differential geometric tools for their analysis. Thus, one is interested in how to extend information geometry. Since equilibrium thermodynamics can be formulated with information geometry, a geometrization of nonequilibrium thermodynamics is one of the keys for this task. Related to this geometrized nonequilibrium thermodynamics, it is of interest to explore how duality emphasized in information geometry appears in nonequilibrium thermodynamics. This duality may be related to other mathematical notions, including adjoint symmetry [7].

Contact geometry is known to be an odd-dimensional counterpart of symplectic geometry [8, 9]. Along with the context stated above, extensions of information geometry are to use contact geometry [10, 11] and para-contact metric contact one [12]. These extensions are consistent with geometrization of equilibrium thermodynamics [13–18]. In [10], a contact geometric description of relaxation processes in nonequilibrium thermodynamic systems was proposed, and a link between information geometry and contact geometry was clarified. Similarly, in [12], para-contact metric contact metric geometric descriptions were proposed. Since Markov chains can be used as models of nonequilibrium thermodynamic systems [19], it can be expected that contact geometry and Markov chains are related. From this, master equations and Markov chain Monte Carlo (MCMC) methods are expected to be formulated in terms of contact and information geometries.

Continuous-time master equations are first order ordinary differential equations. These and Fokker–Plank equations have been used to describe the time-developments of distribution functions [20, 21]. Master equations are closely related to MCMC methods [22] and have been used to model nonequilibrium statistical phenomena [19]. Further development...
of theory of master equations is expected to yield those of MCMC methods and nonequilibrium statistical physics. MCMC methods are well-known methods for obtaining expectation values (averaged values) of some quantities with respect to target distribution functions numerically. These methods have been applied in various disciplines including mathematical engineering, physics, statistics and so on [23–28]. For constructing MCMC methods, master equations with their Markov kernels play fundamental roles. Also it should be noted that dynamical systems on statistical manifolds without contact geometry have been studied in the literature [29–32]. In addition, information geometric descriptions for Markov chains were investigated in the literature as well [33–35].

In this paper a class of dynamical systems describing expectation variables exactly derived from continuous-time master equations is studied from the viewpoints of information geometry and contact geometry, where such master equations consist of a set of appropriately chosen Markov kernels. To formulate this class of dynamical systems with contact geometry, configuration space is identified with a contact manifold, and dynamics is described as contact Hamiltonian dynamical systems. It is then shown that the time-asymptotic limit of the expectation variables defined in closed dynamical systems is consistent with information geometry. Also, convergence rates are explicitly calculated after introducing a metric tensor field. Since the present study is closely related to theorem 4.3 in [10], the main contributions of the present study from the previous ones are stated here. In this paper, (i) a general class of master equations is introduced such that a particular example used in [10] is included, (ii) a nonequilibrium free-energy is introduced, (iii) a convergence rate is calculated, and (iv) duality in the sense of that used in information is stated even in nonequilibrium states.

Some new terminologies will appear in this paper, and they play various roles. The relations among them are briefly listed in table 1.

Table 1. Relations among introduced systems, equations, and methods.

| Sys. for distribution func. | Sys. for observables | Geometric descriptions of observables |
|----------------------------|----------------------|--------------------------------------|
| Primary master Eq.         | Primary moment dynamical Sys. | Contact Hamiltonian Sys. in proposition 4.3 |
| Dual master Eq.            | Dual moment dynamical Sys.     | Contact Hamiltonian Sys. in proposition 4.4 |

2. Solvable master equations

In this section a set of master equations with particular Markov kernels is introduced, and then its solvability is shown first. In arguing this, it is shown how master equations play roles. Then, the idea called ‘dual’ used in information geometry is imported to the master equations. With this idea, dual master equation is introduced. To emphasize this duality, master equation is also referred to as primary master equation.

2.1. Primary master equation

In the following a class of master equations is introduced, and its solvable feature is explicitly clarified. After this, primary master equation is defined.

Let $\Gamma$ be a set of finite discrete states, $t \in \mathbb{R}$ time, and $p(j, t) \, dt$ a probability that a state $j \in \Gamma$ is found in between $t$ and $t + \, dt$. The first objective is to realize a given distribution function $p_0^{\text{eq}}$ that can be written as

$$p_0^{\text{eq}}(j) = \frac{\pi_0(j)}{Z(\theta)},$$

where $\theta \in \Theta \subset \mathbb{R}^n$ is a parameter set with $\theta = \{\theta^1, \ldots, \theta^n\}$, and $Z: \Theta \to \mathbb{R}$ the so-called a partition function so that $p_0^{\text{eq}}$ is normalized. Thus

$$Z(\theta) = \sum_{j \in \Gamma} \pi_0(j), \quad \text{so that} \quad \sum_{j \in \Gamma} p_0^{\text{eq}}(j) = 1.$$

One way to achieve the objective is to employ a special set of Markov kernels. In what follows, how this objective is done is shown.

Let $p: \Gamma \times \mathbb{R} \to \mathbb{R}_{\geq 0}$ be a time-dependent probability function. Then, consider the set of master equations

$$\frac{\partial}{\partial t} p(j, t) = \sum_{j' \in \Gamma} [w(j,j')p(j', t) - w(j'|j)p(j, t)], \quad (1)$$

where $w: \Gamma \times \Gamma \to I, \ (I = \{0, 1\} \subset \mathbb{R})$ is such that $w(j,j') \in I$ denotes a probability that a state jumps from $j'$ to $j$.

Assume the following.

- Choose $w$ to be

$$w_0(j,j') = p_0^{\text{eq}}(j).$$

- The target distribution $p_0^{\text{eq}}(j)$ for any state $j \in \Gamma$ does not vanish: $p_0^{\text{eq}}(j) \neq 0$. 

This paper is organized as follows. In section 2, a set of master equations is introduced by choosing Markov kernels, and such equations are termed primary master equations. After dual master equations are introduced, the explicit solutions of these equations are shown. In section 3, with the use of the clarified features of the primary and dual master equations, differential equations describing time-development of some observables are derived. In section 4, a geometrization of discussions in section 3 is given. Finally, section 5 summarizes this paper and discusses some future works.
To show an explicit form of $p(j, t)$, one rewrites (1) with the assumptions. The equation of motion for $p$ is derived from
\[
\sum_{j'} p(j', t) = 1 - p(j, t),
\]
as
\[
\frac{\partial}{\partial t} p(j, t) = \sum_{j'} [w_0(j) - w_0(j')p(j', t)] - w_0(j)p(j, t)
\]
\[
= p^\text{eq}_0(j) - p(j, t). \tag{2}
\]

In this paper (2) is termed.

**Definition 2.1 (Primary master equations).** The set of equations (2) is referred to as the set of primary master equations.

Then the following proposition holds.

**Proposition 2.1 (Solutions of the primary master equations).** The solution of (2) is
\[
p(j, t) = e^{-t} p(j, 0) + (1 - e^{-t}) p^\text{eq}_0(j). \tag{3}
\]
Then it follows that
\[
\lim_{t \to \infty} p(j, t) = p^\text{eq}_0(j). \tag{4}
\]

**Proof.** Solving (2) for $p(j, t)$ with a set of given initial values $\{p(j, 0)\}$, one has the explicit form of $p(j, t)$, (3). From this explicit form, the long-time limit of $p(j, t)$ can be given immediately as (4).

With this proposition, one notices the following.

1. Any solution $p$ depends on $\theta$.
2. Any state transition from $j$ to $j'(xj)$ does not occur with this set of master equations.
3. If $\sum_j p(j, 0; \theta) = 1$, then $\sum_j p(j, t; \theta) = 1$ for any $t > 0$.

Taking into account 1 above, $p(j, t)$ is denoted $p(j, t; \theta)$.

### 2.2. Dual master equation

In this subsection, some features of target distribution functions $p^\text{eq}_0$ are discussed first. These provide a motivation for introducing the dual of the introduced master equations. Such master equations and dual master equations will be used in section 3 to derive dynamical systems for observables.

Target distribution functions considered in this paper are the exponential family, since they are discrete distribution functions. This family is a class of probability functions defined as follows. Let $\theta = \{\theta^a\} \in \Theta \subseteq \mathbb{R}^\alpha$ be a finite parameter set, and $x$ a set of random variables. If a probability function $p_\theta$ being parameterized by $\theta$ can be written of the form
\[
p_\theta(x) = \exp(\theta^a \xi_a(x) - \Psi(\theta) + C(\xi)), \tag{5}
\]
with some functions $\{\xi_a\}$, $C$ and $\Psi$, then $p_\theta$ is said to belong to the exponential family. In this paper, the Einstein convention, when an index variable appears twice in a single term it implies summation of all the values of the index, is adapted. In what follows the function $C$ is assumed to be eliminated. Also, the following is assumed.

**Postulate 2.1.** In (5), $\theta^a$ and $E[\xi_a]$ for each $a$ form a pair of dimensionless thermodynamic conjugate variables, where $E[\xi_a]$ is the expectation value of $\xi_a$ with respect to $p_\theta$.

The function $\Psi$ is used for normalizing $p_\theta$, and gives the quantities
\[
\eta_a = E[\xi_a] = \frac{\partial \Psi}{\partial \theta^a}, \quad a \in \{1, \ldots, n\} \tag{6}
\]
uniquely. Then the correspondence between $\theta^a$ and $\eta_a$ is one-to-one. This uniqueness provides a motivation for introducing the other expression of $p^\text{eq}_0$. Substituting $\eta = \eta(\theta)$ into $p^\text{eq}_0$, one has the other expression of the target distribution. This is denoted by $p^\eta_0$, and the corresponding master equations are obtained from (2) as
\[
\frac{\partial}{\partial t} p(j, t) = p^\eta_0(j) - p(j, t). \tag{7}
\]
This equation is termed in this paper as follows.

**Definition 2.2 (Dual master equations).** The set of equations (7) is referred to as the set of dual master equations associated with (2).

The primary master equations are used for the cases where the set $\theta = \{\theta^a\}$ is kept fixed, and $\eta$ depends on random variables. On the other hand, the dual master equations are used for the cases where $\eta = \{\eta_a\}$ in (6) is kept fixed, and $\theta$ depends on random variables.

Similar to proposition 2.1, one has the following.

**Proposition 2.2. (Solutions of the dual master equations) The solution of (7) is
\[
p(j, t) = e^{-t} p(j, 0) + (1 - e^{-t}) p^\eta_0(j). \tag{8}
\]
Then it follows that
\[
\lim_{t \to \infty} p(j, t) = p^\eta_0(j). \tag{9}
\]

**Proof.** A way to prove this is analogous to the proof of proposition 2.1.

Accordingly, the solution of this equation is denoted $p(j, t; \eta)$.

In this paper the primary and the dual master equations, (2) and (7), are referred to as solvable master equations.

Before closing this section, attention is concentrated on a spin system to show how the general theory developed in this section is applied to a physical model.
Example. (Kinetic Ising model without interaction). Let \( \sigma \) be a spin variable that takes the values \( \sigma = \pm 1 \), \( \mathcal{H} \) a constant magnetic field whose dimension is an energy, \( T \) the absolute temperature, and \( \theta = \mathcal{H}/(k_B T) \) with \( k_B \) being the Boltzmann constant. Consider the equilibrium system consisting of the one-spin system being coupled with a heat bath with \( T \). With these introduced variables, the canonical distribution is given by

\[
p_{\text{Ising}}^\text{eq}(\sigma) = \exp(\theta \sigma - \psi_{\text{Ising}}^\text{eq}(\theta)),
\]

where \( Z_{\text{Ising}}(\theta) \) is calculated to be

\[
Z_{\text{Ising}}(\theta) = \sum_{\sigma = \pm 1} \exp(\theta \sigma) = 2 \cosh(\theta).
\]

When this canonical distribution is chosen to be the target distribution function, the set of primary master equations is immediately obtained from (2) as

\[
\frac{\partial}{\partial \theta} p_{\text{Ising}}(\sigma, t; \theta) = \exp(\theta \sigma - \psi_{\text{Ising}}^\text{eq}(\theta)) - p_{\text{Ising}}(\sigma, t; \theta),
\]

\[
\sigma = \pm 1,
\]

where

\[
\psi_{\text{Ising}}^\text{eq}(\theta) := \ln Z_{\text{Ising}}(\theta) = \ln(2 \cosh(\theta)).
\]

Figure 1(a) shows the graph of \( \psi_{\text{Ising}}^\text{eq} \), and (b) shows a schematic picture of the time-development of \( p_{\text{Ising}} \). Since \( p_{\text{Ising}}^\text{eq}(+1, t; \theta) + p_{\text{Ising}}^\text{eq}(-1, t; \theta) = 1 \) holds for any \( t \), the dynamics takes place on this one-dimensional line.

The set of dual master equations is derived as follows.

Finally, the explicit form of (7) is obtained as

\[
\frac{\partial}{\partial \theta} p_{\text{Ising}}(\sigma, t; \eta) = \frac{1 - \eta^2}{2} \exp(\sigma \Tanh^{-1}\eta) - p_{\text{Ising}}(\sigma, t; \eta),
\]

\[
\sigma = \pm 1.
\]

The equilibrium free-energy \( \mathcal{F}_{\text{Ising}}^\text{eq} \) is written in terms of \( \psi_{\text{Ising}}^\text{eq} \) in (11) as

\[
\mathcal{F}_{\text{Ising}}^\text{eq} = -k_B T \psi_{\text{Ising}}^\text{eq},
\]

from which one can interpret \( \psi_{\text{Ising}}^\text{eq} \) as the negative dimensionless free-energy for this model.

This system was briefly studied in [10], and is referred to as the kinetic Ising model without interaction in this paper. Note that this system is simplified one originally considered in [19].

3. Time-development of observables

In this section differential equations describing time-development of observables are derived with the solvable master equations under some assumptions. Here observable in this paper is defined as a function that does not depend on a random variable or a state. Thus expectation values with respect to a probability distribution function are observables. Then, time-asymptotic limits of such observables are stated.

3.1. Expectation values associated with \( p(j, t; \theta) \)

In this subsection, the case where \( \mathcal{O}_v \) in (5) depends on random variables is considered. First, one defines some expectation values.

Definition 3.1 (Expectation value with respect to \( p(j, t; \theta) \)). Let \( \mathcal{O}_v : \Gamma \to \mathbb{R} \) be a function with \( a \in [1, \ldots, \mathcal{A}] \), and \( p : \Gamma \times \mathbb{R} \to [0,1] \) a distribution function that follows (2). Then

\[
\langle \mathcal{O}_v \rangle_p(t) := \sum_{j \in \Gamma} \mathcal{O}_v(j)p(j, t; \eta),
\]

and

\[
\langle \mathcal{O}_v \rangle^\text{eq} := \sum_{j \in \Gamma} \mathcal{O}_v(j)p^\text{eq}(j),
\]
are referred to as the expectation value of $C_i$ with respect to $p$, and that with respect to $p^\text{eq}_i$, respectively.

If a target distribution function belongs to the exponential family, then the function $\Psi^\text{eq} : \Theta \to \mathbb{R}$ with

$$\Psi^\text{eq}(\theta) := \ln \left( \sum_{j \in \Gamma} e^{\theta \, C_i(j)} \right),$$  \hspace{1cm} \text{(13)}

plays various roles. Here and in what follows, \(13\) is assumed to exist. In the context of information geometry, this function is referred to as the $\theta$-potential. Discrete distribution functions are considered in this paper and it has been known that such distribution functions belong to the exponential family, then $\Psi^\text{eq}$ in \(13\) also plays a role throughout this paper. The value of the function $\Psi^\text{eq}(\theta)$ can be interpreted as the negative dimension-less free-energy (see \(12\)).

One then can extend this function to that for nonequilibrium states as $\Psi : \Theta \times \mathbb{R} \to \mathbb{R}$ with

$$\Psi(\theta, t) := \left( \int_0^1 \sum_{j \in \Gamma} p(j, t; \theta) \right) \Psi^\text{eq}(\theta),$$  \hspace{1cm} \text{where \(J^0 := \sum_j 1\)} \hspace{1cm} \text{(14)}

Since $p^\text{eq}_i(j) = 0$ and $\Psi^\text{eq}(\theta) < \infty$ by assumptions, the function $\Psi$ \(14\) exists. Extending the idea for the equilibrium case, the function $\Psi$ may be interpreted as a nonequilibrium negative dimension-less free-energy. Although there could be more suitable free-energy for nonequilibrium states, \(14\) is employed in this paper.

One notices the following.

Remark 3.1.

1. Relations among the expectation values of quantities introduced above in the time-asymptotic limit are found with \(4\) as

$$\lim_{t \to \infty} \langle C_i \rangle_\theta(t) = \langle C_i \rangle^\text{eq}_\theta, \text{ and } \lim_{t \to \infty} \Psi(\theta, t) = \Psi^\text{eq}(\theta),$$

2. Since a target distribution function belongs to the exponential family, it follows that \[1\]

$$\ln p^\text{eq}_i \theta_\theta = \mathbb{L}[\Psi^\text{eq}](\eta), \text{ with } \eta_\alpha = \frac{\partial \Psi^\text{eq}(\theta)}{\partial \theta^\alpha}. \hspace{1cm} \text{(15)}$$

Here $\mathbb{L}[\Psi^\text{eq}]$ is the Legendre transform of $\Psi^\text{eq}$,

$$\mathbb{L}[\Psi^\text{eq}](\eta) := \sup_{\theta \in \Theta} \left[ \theta^\alpha \eta_\alpha - \Psi^\text{eq}(\theta) \right].$$

3. Since a target distribution function is a discrete one and thus belongs to the exponential family, one has (see \(6\) and \[1\])

$$\langle C_i \rangle^\text{eq}_\theta = \frac{\partial \Psi^\text{eq}(\theta)}{\partial \theta^\alpha}. \hspace{1cm} \text{(16)}$$

Also, define cross entropy $H$, and negative entropy at equilibrium $H^\text{eq}$ so that

$$H(\theta, t) := \sum_{j \in \Gamma} p(j, t; \theta) \ln p^\text{eq}_i \theta \, (j) = \langle \ln p^\text{eq}_i \theta \rangle_\theta, \text{ and }$$

$$H^\text{eq}(\theta) := \sum_{j \in \Gamma} p^\text{eq}_i \theta(p^\text{eq}_i \theta(j)) = \langle \ln p^\text{eq}_i \theta \rangle^\text{eq}_\theta. \hspace{1cm} \text{(17)}$$

From these definitions with \(4\), the time-asymptotic limit of $H$ is obtained as

$$\lim_{t \to \infty} H(\theta, t) = H^\text{eq}(\theta).$$

3.2. Expectation values associated with $p(j, t; \eta)$

From discussions on dual master equations, it is natural to consider systems where $\theta = \{\theta^\alpha\}$ depends on random variables. For this purpose, the set $\theta$ is treated as variables depending on $j$ in this subsection.

One defines the following.

Definition 3.2 (Expectation value with respect to $p(j, t; \eta)$). Let $\theta^\alpha : \Gamma \to \mathbb{R}$ be a function with $a \in \{1, \ldots, n\}$, and $p : \Gamma \times \mathbb{R} \to \mathbb{R}_{\geq 0}$ a distribution function that follows \(7\). Then

$$\langle \theta^\alpha \rangle_\eta(t) := \sum_{j \in \Gamma} \theta^\alpha(j) p(j, t; \eta), \text{ and }$$

$$\langle \theta^\alpha \rangle^\text{eq}_\eta := \sum_{j \in \Gamma} \theta^\alpha(j) p^\text{eq}_i \theta_\theta(j),$$

are referred to as the expectation value of $\theta^\alpha$ with respect to $p$, and that with respect to $p^\text{eq}_i \theta$, respectively.

One notices the following.

Remark 3.2.

1. Relations among the expectation values of quantities introduced above in the time-asymptotic limit are found with \(9\) as

$$\lim_{t \to \infty} \langle \theta^\alpha \rangle_\eta(t) = \langle \theta^\alpha \rangle^\text{eq}_\eta,$$

2. The following relation is satisfied with the set of time-independent variables $\eta = \{\eta_1, \ldots, \eta_n\}$

$$\mathbb{L}[\Psi^\text{eq}](\eta) = H^\text{eq}(\theta(\eta)), \text{ with } \eta_\alpha = \frac{\partial \Psi^\text{eq}(\theta)}{\partial \theta^\alpha}. \hspace{1cm} \text{(18)}$$

Here the function $H^\text{eq}$ as a function of $\eta$ is obtained as the following two steps. (i) Solving $\eta_\alpha = \partial \Psi^\text{eq}(\theta)/\partial \theta^\alpha$ for $\theta^\alpha$, one writes $\theta^\alpha = \theta^\alpha(\eta)$. (ii) Substituting $\theta^\alpha = \theta^\alpha(\eta)$ into $H^\text{eq}(\theta)$, one has $H^\text{eq}(\theta(\eta))$. Letting $\Psi^\text{eq}$ be the Legendre transform of $\Psi^\text{eq}$, one can write (18) as

$$\Psi^\text{eq}(\eta) = H^\text{eq}(\theta(\eta)).$$
In the context of information geometry, the function $\Phi_{eq}$ is referred to as the $\eta$-potential.

3. It can be shown that [1]

$$
\langle \theta^a \rangle_{eq} = \frac{\partial \Phi_{eq}}{\partial \eta_a},
$$

(19)

and

$$
\frac{\partial^2 \Phi_{eq}}{\partial \theta^a \partial \theta^b} = \delta^a_b,
$$

(20)

where $\delta^a_b$ is the Kronecker delta giving unity for $a = b$, and zero otherwise.

Similar to (17), one defines

$$
H(\eta, t) = \sum_{j \in \Gamma} p(j, t; \eta) \ln p^eq(j) = \langle \ln p^eq \rangle_{\eta},
$$

and

$$
H^{eq}(\eta) = \sum_{j \in \Gamma} p^eq(j) \ln p^eq(j) = \langle \ln p^eq \rangle_{\eta}.\]

By definition, it follows that $H^{eq}(\theta(\eta)) = H^{eq}(\eta)$, where $\theta(\eta)$ is obtained by solving $\eta^b = \partial \Phi_{eq} / \partial \theta^b$ for $\theta^a$.

Explicit forms of the expectation values and ones of remarks 3.1 and 3.2 for example in section 2 can be shown as follows.

**Example.** Consider the kinetic Ising model without interaction introduced in example in section 2. Choose $\mathcal{O}$ as

$$
\mathcal{O}(\sigma) = \sigma,
$$

from which one has the expectation value of $\mathcal{O}$ with respect to $p^eq$ as

$$
\langle \sigma \rangle_{eq} = \sum_{\sigma = \pm 1} \sigma p^eq(\sigma) = e^\theta - e^{-\theta} / Z_{\text{Ising}}(\theta) = \tanh(\theta).
$$

The $\theta$-potential $\Psi_{eq}$ defined in (13) is calculated to be

$$
\Psi_{\text{Ising}}^{eq}(\theta) = \ln \left( \sum_{\sigma = \pm 1} e^{\theta \sigma} \right) = \ln (2 \cosh(\theta)) = \psi_{\text{Ising}}^{eq}(\theta),
$$

where $\psi_{\text{Ising}}^{eq}(\theta)$ has been defined in (11). Then, one has $\eta$ in (15) as

$$
\eta = \frac{d\Psi_{\text{Ising}}^{eq}(\theta)}{d\theta} = \tanh(\theta).
$$

(21)

Also it turns out that $\Psi_{\text{Ising}}^{eq}$ is strictly convex due to

$$
\frac{d^2\Psi_{\text{Ising}}^{eq}(\theta)}{d\theta^2} = \frac{d\eta}{d\theta} = \text{sech}^2(\theta) > 0.
$$

(22)

The negative entropy at equilibrium $H^{eq}_{\text{Ising}}$ is calculated to be

$$
H^{eq}_{\text{Ising}}(\theta) = \sum_{\sigma = \pm 1} p^eq(\sigma) \ln p^eq(\sigma)
= \sum_{\sigma = \pm 1} p^eq(\sigma) (\theta \sigma - \psi_{\text{Ising}}(\theta))
= \theta \tanh(\theta) - \ln (2 \cosh(\theta)).
$$

(23)

The relation (18) is verified for this example below. First, the Legendre transform $\mathbb{L} [\Psi_{\text{Ising}}^{eq}(\eta)]$ is calculated as

$$
\mathbb{L} [\Psi_{\text{Ising}}^{eq}(\eta)] = \sup_{\theta} [\theta \eta - \Psi_{\text{Ising}}^{eq}(\theta)]
= [\theta \eta - \Psi_{\text{Ising}}^{eq}(\theta)]_{\theta = \tanh^{-1}(\eta)}
= \eta \tanh^{-1}(\eta) - \ln [2 \cosh(\tanh^{-1}(\eta))].
$$

Second, it follows from (21) and (23) that $\mathbb{L} [\Psi_{\text{Ising}}^{eq}(\eta)] = H^{eq}_{\text{Ising}}(\theta(\eta))$. They are denoted by $\Phi_{\text{Ising}}^{eq}(\eta) = \mathbb{L} [\Psi_{\text{Ising}}^{eq}(\eta)] = H^{eq}_{\text{Ising}}(\theta(\eta))$. Another expression of $\Phi_{\text{Ising}}^{eq}(\eta)$ is obtained from

$$
\tanh^{-1}(x) = \frac{1}{2} \ln \left( \frac{1 + x}{1 - x} \right),
$$

and

$$
\tanh^{-1}(x) = \cosh^{-1} \left( \frac{1}{\sqrt{1 - x^2}} \right), \quad |x| < 1,
$$

as

$$
\Phi_{\text{Ising}}^{eq}(\eta) = \mathbb{L} [\Psi_{\text{Ising}}^{eq}(\eta)] = \frac{\eta}{2} \ln \left( \frac{1 + \eta}{1 - \eta} \right)
= \frac{1}{2} \ln(1 - \eta^2) - \ln 2, \quad |\eta| < 1.
$$

where $-1 < \eta < 1$ coming from (21) has been applied. Similar to $\Phi_{\text{Ising}}^{eq}$, the function $\Phi_{\text{Ising}}^{eq}$ is shown to be strictly convex directly. It follows from

$$
\frac{dH^{eq}_{\text{Ising}}}{d\theta} = \theta \text{sech}^2(\theta),
$$

and

$$
\frac{dH^{eq}_{\text{Ising}}}{d\eta} = \frac{d\Phi_{\text{Ising}}^{eq}}{d\eta} = \frac{d\Phi_{\text{Ising}}^{eq}}{d\theta} = \theta(\eta),
$$

that

$$
\frac{d^2\Phi_{\text{Ising}}^{eq}}{d\eta^2} = \frac{d\theta}{d\eta} = \cosh^2[\theta(\eta)]
= \cosh^2[\tanh^{-1}(\eta)] = \frac{1}{1 - \eta^2} > 0.
$$

(24)

The relation (20) for this example is verified by combining (22) and (24) as

$$
\frac{d^3\Phi_{\text{Ising}}^{eq}}{d\eta^2} = \frac{d^2\Phi_{\text{Ising}}^{eq}}{d\eta^2} = 1.
$$

Before closing this subsection, physical interpretations of $\theta^a$ and $\langle \theta^a \rangle_{eq}$ are argued by taking a simple example. Identify $\theta = \mathcal{H} / (k_B T)$, where $\mathcal{H}$ is a fixed constant magnetic field and $T$ a temperature of a fixed heat bath. Then, $\theta$ is roughly speaking the inverse temperature of the heat bath. It should be emphasized that $T$ does not depend on time, and thus it is constant. On the other hand, one can introduce another temperature $\langle T \rangle'$ that is proportional to the expectation value of the squared velocity of a particle $\langle v^2 \rangle'$, where $\langle \cdots \rangle'$ has been introduced appropriately. In a gas case, such a relation is $\langle T \rangle' = m \langle v^2 \rangle' / (3 k_B)$, where $m$ is the mass of a particle.

This leads to $\langle \theta \rangle' = \mathcal{H} / (k_B \langle T \rangle')$, which is not constant.
in time in general. Although \( \langle \cdots \rangle \) is not same as \( \langle \cdots \rangle_0 \), one can generalize this example to the case where \( \{ \theta^a \} \) is a fixed parameter set and \( \{ \langle \theta^a \rangle \} \) an averaged one.

### 3.3. Closed dynamical systems

As shown below a set of differential equations for \( \{ \langle Q_a \rangle \} \) and \( \Psi \) can be found in a closed form. When expectation values depending on time are seen as variables that satisfy differential equations, they are referred to as expectation variables in this paper.

**Proposition 3.1 (Dynamical system obtained from the primary master equations).** Let \( \theta \) be a time-independent parameter set characterizing a discrete target distribution function \( p^\text{eq}_q \). Then \( \{ \langle Q_a \rangle \} \) and \( \Psi \) are solutions to the differential equations on \( \mathbb{R}^{2n+1} \)

\[
\frac{d}{dt} \theta^a = 0, \quad \frac{d}{dt} \{ \langle Q_a \rangle \} = - \{ \langle Q_a \rangle \} + \frac{\partial \Psi^\text{eq}}{\partial \theta^a}, \quad \text{and} \quad \frac{d}{dt} \Psi = -\Psi + \Psi^\text{eq}.
\]

**Proof.** Nontrivial parts of this proof are for the second and third differential equations. A proof that the equation for \( \{ \langle Q_a \rangle \} \) holds is given first. It follows from (2) and (16) that

\[
\frac{d}{dt} \{ \langle Q_a \rangle \} = \sum_{j \in \Gamma} \mathcal{O}_a(j) \frac{\partial p(j, t; \theta)}{\partial t} = \sum_{j \in \Gamma} \mathcal{O}_a(j)[p^\text{eq}_q(j) - p(j, t; \theta)] = - \{ \langle Q_a \rangle \} + \frac{\partial \Psi^\text{eq}}{\partial \theta^a}.
\]

Then, a proof that the equation for \( \Psi \) holds is given below. It follows from (2) and (14) that

\[
\frac{d}{dt} \Psi = \left( \frac{1}{\beta} \sum_{j \in \Gamma} \frac{1}{p^\text{eq}_q(j')} \frac{\partial p(j', t; \theta)}{\partial t} \right) \Psi^\text{eq} = \frac{1}{\beta} \sum_{j \in \Gamma} \frac{p^\text{eq}_q(j) - p(j', t; \theta)}{p^\text{eq}_q(j')} \Psi^\text{eq} = -\Psi + \Psi^\text{eq}.
\]

**Remark 3.3.** The explicit time-dependence for this system is obtained as \( \theta^a(t) = \theta^a(0) \)

\[
\{ \langle Q_a \rangle \}_0(t) = e^{-t} \left[ \{ \langle Q_a \rangle \}_0(0) - \frac{\partial \Psi^\text{eq}}{\partial \theta^a} \right] + \frac{\partial \Psi^\text{eq}}{\partial \theta^a}, \quad \text{and} \quad \Psi(\theta, t) = e^{-t}[\Psi(0) - \Psi^\text{eq}(\theta)] + \Psi^\text{eq}(\theta).
\]

From these and (16), one can also verify that the time-asymptotic limit of these variables are those defined at equilibrium:

\[
\lim_{t \to -\infty} \{ \langle Q_a \rangle \}_0(t) = \{ \langle Q_a \rangle \}^\text{eq}_0, \quad \text{and} \quad \lim_{t \to -\infty} \Psi(\theta, t) = \Psi^\text{eq}(\theta).
\]

For the later convenience, this dynamical system is termed follows.

**Definition 3.3 (Primary moment dynamical system).** The dynamical system in proposition 3.1 is referred to as a (primary) moment dynamical system.

Similar to the primary moment dynamical system, one has the following dynamical system.

**Proposition 3.2 (Dynamical system obtained from the dual master equations).** Let \( \{ \eta_a \} \) be a set of time-independent variables satisfying the second equation of (15). Then \( \{ \theta^a \}_0(t) \) and \( H(\eta, t) \) are solutions to the differential equations on \( \mathbb{R}^{2n+1} \)

\[
\frac{d}{dt} \{ \theta^a \}_0 = -\{ \theta^a \}_0 + \frac{\partial \Phi^\text{eq}}{\partial \eta_a}, \quad \frac{d}{dt} \eta_a = 0, \quad \text{and} \quad \frac{d}{dt} H = -H + \Phi^\text{eq}(\eta).
\]

**Proof.** Nontrivial parts of this proof are for the first and third differential equations. A proof that the equation for \( \{ \theta^a \}_0 \) holds is given first. It follows from (7) and (19) that

\[
\frac{d}{dt} \{ \theta^a \}_0 = \sum_{j \in \Gamma} \theta^a(j)[p^\text{eq}_\eta(j) - p(j, t; \eta)] = -\{ \theta^a \}_0 + \frac{\partial \Phi^\text{eq}}{\partial \eta_a}.
\]

Then, a proof that the equation for \( H(\eta, t) \) holds is given as

\[
\frac{dH}{dt} = \sum_{j \in \Gamma} \frac{\partial p(j, t; \eta)}{\partial t} \ln p^\text{eq}_\eta(j) = \sum_{j \in \Gamma} [p^\text{eq}_\eta(j) - p(j, t; \eta)] \ln p^\text{eq}_\eta(j) = -H + \Phi^\text{eq}(\eta).
\]

**Remark 3.4.** The explicit time-dependence for this system is obtained as \( \eta_a(t) = \eta_a(0) \)

\[
\{ \theta^a \}_0(t) = e^{-t} \left[ \{ \theta^a \}_0(0) - \frac{\partial \Phi^\text{eq}}{\partial \eta_a} \right] + \frac{\partial \Phi^\text{eq}}{\partial \eta_a}, \quad \text{and} \quad H(\eta, t) = e^{-t}[H(0) - \Phi^\text{eq}(\eta)] + \Phi^\text{eq}(\eta).
\]

From these and (19), one can also verify that the time-asymptotic limit of these variables are those defined at equilibrium:

\[
\lim_{t \to -\infty} \{ \theta^a \}_0(t) = \{ \theta^a \}^\text{eq}_\eta, \quad \text{and} \quad \lim_{t \to -\infty} H(\eta, t) = \Phi^\text{eq}(\eta).
\]

For the later convenience, this dynamical system is termed follows.

**Definition 3.4 (Dual moment dynamical system).** The dynamical system in proposition 3.2 is referred to as a dual moment dynamical system.
In the following, it is shown how the general theory developed above is applied by focusing on the system stated in example in section 2.

Example. Consider the kinetic Ising model without interaction introduced in example in section 2. Choose $\mathcal{O}$ to be $\mathcal{O}(\sigma) = \sigma$, from which one has the expectation value of $\sigma$ with respect to $p^\mathcal{O}$ as $\langle \sigma \rangle^\mathcal{O} = \tan(\theta)$. The primary moment dynamical system discussed in proposition 3.1 for this case is

$$\frac{d\theta}{dt} = 0, \quad \frac{d\langle \sigma \rangle}{dt} = -\langle \sigma \rangle + \tan(\theta),$$

and

$$\frac{d\Psi_{\text{Ising}}}{dt} = -\Psi_{\text{Ising}} + \ln(2 \cosh(\theta)).$$

The dual moment dynamical system discussed in proposition 3.2 for this case is

$$\frac{d\langle \theta \rangle}{dt} = -\langle \theta \rangle + \tanh^{-1}\eta, \quad \frac{d\eta}{dt} = 0,$$

and

$$\frac{dH_{\text{Ising}}}{dt} = -H_{\text{Ising}} + \eta \tan^{-1}\eta - \ln[2 \cosh(\tan^{-1}\eta)].$$

Since $\theta$ is chosen as $\theta = \mathcal{H}/(k_BT)$ for this example, the equations above correspond to the case where temperature depends on time, $t$, if $\mathcal{H}$ is constant.

4. Geometric description of dynamical systems

For the model used in this paper, the equilibrium states can be geometrized with information geometry, since the target distribution functions belong to the exponential family [1]. Several geometrization of nonequilibrium states for some models and methods have been proposed [33–35]. Yet, sufficient to say that there remains no general consensus on how best to extend information geometry of equilibrium states to a geometry expressing nonequilibrium states. In this section, a geometrization of nonequilibrium states is proposed for the solvable master equations.

Also it is shown that such a geometry is consistent with information geometry in the sense that the time-asymptotic limit of solutions to the primary and dual master equations is described in information geometry. Here one notices that manifolds and geometric framework expressing equilibrium states should be obtained in the limiting case of nonequilibrium ones. Thus, the geometry for expressing nonequilibrium states should be wider than the geometry expressing equilibrium states in some sense. One such a wider geometry is contact geometry [10, 15].

4.1. Geometry of equilibrium states

In what follows the relation between contact geometry and information geometry is briefly reviewed.

Let $\mathcal{C}$ be a $(2n + 1)$-dimensional manifold, $(n = 1, 2, \ldots)$. If a one-form $\lambda$ on $\mathcal{C}$ is provided and satisfies

$$\lambda \wedge d\lambda \wedge \cdots \wedge d\lambda = 0,$$

then the pair $(\mathcal{C}, \lambda)$ is referred to as a contact manifold, and $\lambda$ a contact one-form. It has been known that there exists a special set of coordinates $(x, y, z)$ with $x = [x^1, \ldots, x^n]$ and $y = [y_1, \ldots, y_n]$ such that $\lambda = dz - y_i dx^i$. The existence of such coordinates is guaranteed mathematically stated as the Darboux theorem [9]. The Legendre submanifold $\mathcal{A} \subset \mathcal{C}$ is an $n$-dimensional manifold where $\lambda|_{\mathcal{A}} = 0$ holds. One can verify that

$$\mathcal{A}_{\mathcal{C}_{\mathcal{A}}} = \left\{ (x, y, z) \mid y_i = \frac{\partial\varpi}{\partial x^i}, \text{ and } z = \varpi(x) \right\}.$$ (26)

is a Legendre submanifold, where $\varpi : \mathcal{C} \rightarrow \mathbb{R}$ is a function on $\mathcal{C}$. The submanifold $\mathcal{A}_{\mathcal{C}_{\mathcal{A}}}$ is referred to as the Legendre submanifold generated by $\varpi$, and is used for describing equilibrium thermodynamic systems [14]. It should be noted that how statistical mechanics dealing with distribution functions adopts Legendre submanifolds was investigated [15, 36].

Equilibrium states, or equivalently target states, are identified with the Legendre submanifolds generated by functions in the context of geometric thermodynamics [13, 14]. Besides, in the context of information geometry, equilibrium states are identified with dually flat spaces. Combining these identifications, one has the following.

Proposition 4.1 (A contact manifold and a strictly convex function induce a dually flat space, [10]). Let $(\mathcal{C}, \lambda)$ be a contact manifold, $(x, y, z)$ a set of coordinates such that $\lambda = dz - y_i dx^i$ with $x = [x^1, \ldots, x^n]$ and $y = [y_1, \ldots, y_n]$, and $\varpi$ a strictly convex function depending only on $x$. If the Legendre submanifold generated by $\varpi$ is simply connected, then $(\mathcal{C}, \lambda, \varpi)$ induces the n-dimensional dually flat space

Dually flat space is defined in information geometry, and this space consists of a manifold $\mathcal{M}$, a (pseudo-) Riemannian metric tensor field $g^F$, and flat connections $\nabla$, $\nabla^F$ that satisfy

$$X[g^F(Y, Z)] = g^F(\nabla_Y Z) + g^F(Y, \nabla^F_Z Z), \quad \forall X, Y, Z \in T\mathcal{M}.$$ (27)

Thus a dually flat space is a quadruplet $(\mathcal{M}, g^F, \nabla, \nabla^F)$. To apply this proposition to physical systems, the coordinate sets $x$ and $y$ are chosen such that $x^a$ and $y_i$ form a thermodynamic conjugate pair for each $a$. Here it is assumed that such thermodynamic variables can be defined even for nonequilibrium states, and they are consistent with those variables defined at equilibrium. In addition to this, the physical dimension of $\varpi$ should be equal to that of $y_i dx^i$.

How to specify a function $\varpi$ in proposition is given as follows. In mathematical statistics, given a probability distribution function $p_0$ parameterized by $\theta$, the Fisher information matrix is defined. Each component is defined by

$$g^F_{\theta \theta}(\theta) = \mathbb{E} \left[ \frac{\partial \ln p_0}{\partial \theta^a} \frac{\partial \ln p_0}{\partial \theta^b} \right].$$ (28)

For the exponential family, these matrix components can be obtained not only by calculating (28), but also by differentiating $\Psi^{eq}$ in (13) with respect to $\theta$ twice [1]. This $\Psi^{eq}$ and its Legendre transform $\Phi^{eq} = L[\Psi^{eq}]$, a $\theta$-potential and an
η-potential, are chosen as ω in this proposition. This choice is a part of procedure linking the space of distribution functions and a contact manifold.

Instead of discussing physical meaning of the proposition, we explain how this proposition applies to a physical system by analyzing example in section 2.

**Example.** Consider the kinetic Ising model without interaction introduced in example in section 2. Let C be the three-dimensional manifold with $C \cong \mathbb{R} \times (-1, 1) \times \mathbb{R}$, and $\lambda = x - y \, dx$. Identify the coordinates such that $x = \theta$, $y = (\sigma \, )_0$, and $z = \Psi$, then the condition (25) is $d\theta \wedge d(\sigma \, )_0 \wedge d\Psi \equiv 0$. Then the Legendre submanifold generated by $\Psi_{\text{Ising}}$,

$$A_{\Psi_{\text{Ising}}} = \{(\theta, (\sigma \, )_0, \Psi) \mid (\sigma \, )_0 \} = \frac{d\Psi_{\text{Ising}}}{d\theta} \text{ and } \Psi = \Psi_{\text{Ising}}(\theta) \}, \quad (29)$$

is a suitable submanifold expressing the equilibrium state. It is known that equilibrium thermodynamic systems can be described in terms of information geometry. For example, the Fisher metric tensor field and the expectation coordinate on $A_{\Psi_{\text{Ising}}}$ are

$$g_{\text{Ising}} = \frac{F}{\Psi_{\text{Ising},0}} \, d\theta \otimes d\theta = \text{sech}^2 \theta \, d\theta \otimes d\theta,$$

$$\eta = \frac{d\Psi_{\text{Ising}}}{d\theta} = \tanh(\theta),$$

respectively, where (22), (21), and

$$g_{\text{Ising,0}} = \frac{d^2\Psi_{\text{Ising}}}{d\theta^2},$$

have been used. Combining these and the first equality of (24), one verifies that $\theta$ and $\eta$ are dual with respect to $g_{\text{Ising}}$ in the sense that [1]

$$g_{\text{Ising}} \left( \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \eta} \right) = \left( \text{sech}^2 \theta \right) \frac{d\theta}{d\eta} = 1.$$

Since the Fisher metric tensor filed is a type of Riemannian metric tensor field and it is known that a Riemannian metric tensor field induces the Levi-Civita connection uniquely [37], $g_{\text{Ising}}$ induces the Levi-Civita connection $\nabla^{\text{Ising}}$. Its connection component $\Gamma^{(i)0}_{00}$ is such that $\nabla^{\text{Ising}}_\partial \theta = \Gamma^{(i)0}_{00} \partial_i$, $(\partial \equiv \partial/\partial \theta)$, and its explicit form is calculated as [37]

$$\Gamma^{(i)0}_{00} = \frac{1}{2} \frac{dF}{d\theta} \frac{dF}{d\theta} g_{\text{Ising,0}} = -\text{sech}^2 \theta \, \tanh \theta, \quad \text{(no sum)}$$

where $g_{\text{Ising}} = (g_{\text{Ising,0}})^{-1} = \cosh^2 \theta$. Then one can show

$$\Gamma^{(i)0}_{00} = \frac{1}{2} \frac{dF}{d\theta} g_{\text{Ising,0}} = -\text{sech}^2 \theta \, \tanh \theta, \quad \text{(no sum)}$$

where $C_{\text{Ising,0}}$ is the component of the cubic-form, and is related to the Levi-Civita connection. The cubic-form is defined in information geometry as $C^F = \nabla g^F$ with $\nabla$ being a connection [38]. If $V$ is flat, then its component expression for this example is obtained as

$$C_{\text{Ising}}^F = C_{\text{Ising,0}}(d\theta \otimes d\theta \otimes d\theta), \quad \text{where}$$

$$C_{\text{Ising,0}} = \frac{d^2\Psi_{\text{Ising}}}{d\theta^2} = -2 \, \text{sech}^2 \theta \, \tanh \theta.$$

The $\alpha$-connection defined in information geometry gives dual connections. Since the target distribution function belongs to the exponential family, one can show that the component of the $\alpha$-connection $\Gamma^{(a)00}_{00}$ can be obtained from $\Gamma^{(a)00}_{00} = \frac{1}{2} \frac{dF}{d\theta} \frac{dF}{d\theta} + \frac{1}{2} C_{\text{Ising,0}}$ and

$$\Gamma^{(a)00}_{00} = \frac{1}{2} - \alpha \, C_{\text{Ising,0}}.$$

From this, it follows that $\nabla^{(1)}$ is flat with $\theta$ being its coordinate. Combining the equations above, one has the component expression of (27),

$$\frac{d}{d\theta} \frac{dF}{d\theta} = \Gamma^{(a)00}_{00} + \Gamma^{(a-1)00}_{00}.$$

Furthermore, it can be shown that $\nabla^{(1)}$ is flat with $\eta$ being its coordinate [1]. To summarize, the quadruplet $(A_{\Psi_{\text{Ising}}}^F, g_{\text{Ising}}, \nabla^{(1)}, \nabla^{(1)}_{\text{Ising}})$ is a dually flat space, and is induced from the contact manifold $C$, $\lambda$, and $\Psi_{\text{Ising}}$.

Instead of the identification employed above, one can identify the coordinates such that $x = \eta$, $y = (\theta \, )_0$, and $z = \Phi$ for $C \cong (-1, 1) \times \mathbb{R}^2$. Then, the condition (25) is $d\eta \wedge d(\theta \, )_0 \wedge d\Phi = 0$. In this case the Legendre submanifold generated by $\Phi_{\text{Ising}}^F$,

$$A_{\Phi_{\text{Ising}}} = \{(\eta, (\theta \, )_0, \Phi) \mid (\theta \, )_0 = \frac{d\Phi_{\text{Ising}}}{d\eta}, \text{and } \Phi = \Phi_{\text{Ising}} \},$$

is a suitable submanifold expressing the equilibrium state. On this submanifold, geometric objects such as the Fisher metric tensor field, the cubic form, and the $\alpha$-connection, can explicitly be constructed as well. For example, the Fisher metric tensor field is

$$g_{\text{Ising}} = g_{\text{Ising,0}} \, d\eta \otimes d\eta = \frac{1}{1 - \eta^2} \, d\eta \otimes d\eta,$$

$$\frac{dF}{d\eta} = \frac{d^2\Phi_{\text{Ising}}}{d\eta^2}.$$

The relation between $A_{\Phi_{\text{Ising}}}^F$ and $A_{\Psi_{\text{Ising}}}^F$ can be argued as follows. It follows from (22) and (24) that there exists a diffeomorphism between these submanifolds [10, 15]. Let $\phi_\Phi : A_{\Psi_{\text{Ising}}}^F \rightarrow A_{\Phi_{\text{Ising}}}^F$ be such a diffeomorphism. Then it follows from straightforward calculations [37] that the pull-back of $g_{\text{Ising}}$ is shown to equal to $g_{\text{Ising}}$.

$$\phi_\Phi^*(g_{\text{Ising}}) = g_{\text{Ising}}^F.$$

So far the geometry of equilibrium states has been discussed. One remaining issue is how to give the physical meaning of the set outside $A_{\pm}, \mathcal{C} \setminus A_{\pm}$. A natural
interpretation of $C \setminus A$ would be some set of nonequilibrium states, and is discussed in the next subsection.

### 4.2. Geometry of nonequilibrium states

As shown in proposition 2.1, initial states approach to the equilibrium state as time develops. This can be reformulated on contact manifolds. In the contact geometric framework of nonequilibrium thermodynamics, the equilibrium state is identified with a Legendre submanifold. Then, as it has been clarified in [10, 15], this asymptotic behavior can be identified with a class of contact Hamiltonian vector fields on a contact manifold. Here, contact Hamiltonian vector fields are analogous to symplectic vector fields on symplectic manifolds, and symplectic vector fields correspond to canonical equations of motion in Hamiltonian mechanics. This statement on a class of contact Hamiltonian vector fields can be summarized as follows.

**Proposition 4.2 (Legendre submanifold as an attractor)**

Let $(C, \lambda)$ be a $(2n + 1)$-dimensional contact manifold with $\lambda$ being a contact form, $(x, y, z)$ its coordinates so that $\lambda = dz - y dx^n$ with $x = \{x^1, \ldots, x^n\}$, $y = \{y_1, \ldots, y_n\}$, and $z$ a function depending only on $x$. Then, one has

1. The contact Hamiltonian vector field associated with the contact Hamiltonian $h : C \to \mathbb{R}$ such that
   \[ h(x, y, z) = z \tag{30} \]
   gives
   \[ \frac{d}{dt} x^a = 0, \quad \frac{d}{dt} y^a = \frac{\partial z}{\partial x^a} - y^a, \quad \frac{d}{dt} z = z(x) - z \tag{31} \]
   2. The Legendre submanifold generated by $z$, given by (26), is an invariant manifold for the contact Hamiltonian vector field.
   3. Every point on $C \setminus A$ approaches to $A$ along an integral curve as time develops. Equivalently $A$ is an attractor in $C$.
   4. Let $\{x(0), y(0), z(0)\}$ be a point on $C \setminus A$. Then the relation between the value of $h$ at $t$ and that at $t = 0$ is given by
   \[ h(x(t), y(t), z(t)) = \exp(-t)h(x(0), y(0), z(0)) \tag{32} \]

Applying proposition 4.2 to this system, one has that $\mathcal{A}_\text{eq}$ defined in (29) is an invariant manifold and an attractor. Also, one concludes that $h^\text{eq}_\text{Ising}(\theta, \langle \sigma \rangle_\text{eq}, \Psi)$ decreases exponentially as time develops if an initial point is on $C \setminus \mathcal{A}_\text{eq}$ Ising. From the physical interpretations of $\Psi^\text{eq}_\text{Ising}$ and $\Psi$, one can interpret $h^\text{eq}_\text{Ising}$ in a physical language as follows. The value of $h^\text{eq}_\text{Ising}$ at time $t$ can be interpreted as the difference between dimensionless free-energy at the state specified by $t$ and that at the equilibrium state, realized in the limit $t \to \infty$. This type of contact Hamiltonian will be used in proposition 4.3.

As well as the set of the identifications above, the other set of the identifications $x = \eta, y = \langle \theta \rangle_\eta, z = \Phi$ and $h = h^\Phi_\text{Ising}$ with
\[ h^\Phi_\text{Ising}(\eta, \langle \theta \rangle_\eta, \Phi) = \Phi^\text{eq}_\text{Ising}(\eta) - \Phi \tag{35} \]
enables one to discuss relaxation processes. Then the value of $h^\Phi_\text{Ising}$ at time $t$ can be interpreted as the difference between negative entropy at time $t$ and that at the equilibrium state, realized in the limit $t \to \infty$. This type of contact Hamiltonian will be used in proposition 4.4.

Some basics and physical applications of contact Hamiltonian systems are summarized in [39]. Also for some discussions on nonequilibrium processes in terms of contact geometry, see [15].

Relations among introduced spaces and manifolds are as follows. Let $\mathcal{P}$ and $\mathcal{S}_\theta(\theta \in \Theta \subseteq \mathbb{R}^n) = \{ \eta \}$ be the sets

\[ \mathcal{P} = \{ p \mid p \geq 0 \text{ can be normalized and parameterized by some finite set} \}, \quad \text{and} \quad \mathcal{S}_\theta = \left\{ p \mid p \text{ satisfies (2) and } \sum_{j \in \Gamma} p(j, 0; \theta) = 1 \right\} , \]

respectively. Then the geometrization of $\mathcal{P}$, the space of distribution functions, is information geometry. In particular, dually flat space $(\mathcal{P}, g^F, \nabla, \nabla^*)$ has intensively been studied.
Besides, the present geometry is about a contact manifold \((\mathcal{C}, \lambda)\) provided maps \((\Gamma, \Theta) \to \mathcal{C}\), where these maps consist of the integration of quantities over random variables with the weight of a distribution function in \(\mathcal{S}_0\), and identity maps. Since a target distribution function belongs to the exponential family, a convex function \(\Psi^{eq}\) exists. Then \(((\mathcal{C}, \lambda), \Psi^{eq})\) induces \((\mathcal{A}_{\Psi^{eq}}, \mathcal{S}, \nabla, \nabla^*)\), which enables one to deal with information geometry of equilibrium states from a viewpoint of contact geometry. Also, a map \(\mathcal{C} \to \mathcal{C}\) is realized by a flow of a contact Hamiltonian vector field, and is identified with a nonequilibrium process. As a particular case, choosing a contact Hamiltonian to be \((30)\) with \(\varpi = \Psi^{eq}\), one has a relaxation process. Here such a relaxation process is defined such that an integral curve connects a point of \(\mathcal{C}\), and identity maps.

Remark 4.1. In [10], a class of contact Hamiltonians was chosen for expressing relaxation processes. They are of the form

\[
h(x, y, z) = \tilde{h}(\varpi(x) - z),
\]

with some function \(\tilde{h}\). Note that this class contains \((30)\), which can be verified by choosing \(\tilde{h}\) such that \(\tilde{h}(\Upsilon) = \Upsilon\) with \(\Upsilon = \varpi(x) - z\). In general, the relaxation rate for the case of \((36)\) is not exponential. On the other hand, the relaxation rate for the case of \((30)\) is exponential (See \((32)\)), which is the same as that for \((3)\) and \((8)\). For this reason the original contact Hamiltonian \((30)\) is only considered in this paper.

Remark 4.2. Divergences are defined in information geometry and they play various roles [1, 2]. They are often discussed on dually flat spaces in information geometry. In this paper, connections are not introduced in contact manifolds except for Legendre submanifolds. For this reason aspects on divergences on contact manifolds are not considered here. However it is shown that the negative relative entropy that can be written as a form of divergence \(S_0 \times S_0 \to \mathbb{R}\),

\[
D_{\theta}(p \| p_{\theta}^{eq}) = \sum_{j \in \Gamma} p(j, t; \theta) \ln \left( \frac{p(j, t; \theta)}{p_{\theta}^{eq}(j)} \right),
\]

and the primary master equations yield an inequality for the solvable master equations. Substituting \((2)\) into \((37)\), one has

\[
\frac{d}{dt} D_{\theta}(p \| p_{\theta}^{eq}) = -\sum_{j \in \Gamma} \left( p(j, t; \theta) - p_{\theta}^{eq}(j) \right) \left( \ln p(j, t; \theta) - \ln p_{\theta}^{eq}(j) \right) \leq 0,
\]

where the inequality

\[
(\zeta - \zeta') \left( \ln \zeta - \ln \zeta' \right) \geq 0, \quad \text{for} \quad \zeta, \zeta' > 0
\]

has been used. Similarly one discusses an inequality of the divergence depending on \(p\) and \(p_{\theta}^{eq}\),

\[
D_{\eta}(p \| p_{\eta}^{eq}) = \sum_{j \in \Gamma} p(j, t; \eta) \ln \left( \frac{p(j, t; \eta)}{p_{\eta}^{eq}(j)} \right).
\]

It then follows that

\[
\frac{d}{dt} D_{\eta}(p \| p_{\eta}^{eq}) \leq 0.
\]

Remark 4.3. For nonequilibrium states, it is expected that flows of the solvable master equations form a dually flat space.

4.3. Geometry of primary and dual moment dynamical systems

In this subsection propositions 3.1 and 3.2 are written in a contact geometric language. In what follows phase space is identified with a \((2n + 1)\)-dimensional contact manifold, \((\mathcal{C}, \lambda)\).

As shown below, the dynamical systems stated in these propositions are contact Hamiltonian systems.

Proposition 4.3 (Primary moment dynamical system as a contact Hamiltonian system). The dynamical system in proposition 3.1 can be written as a contact Hamiltonian system.

Proof. Identify \(x, y, z\) and \(\varpi\) in \((31)\) with \(x^a = \Theta^a, \quad y_\eta = \langle \mathcal{O}_\eta \rangle_\theta, \quad \varpi(x) = \Psi^{eq}(\theta), \quad z = \Psi\).

One then sees that \((31)\) is identical to the dynamical system in proposition 3.1. This system is generated by the contact Hamiltonian \(h\) such that

\[
h(\theta, \langle \mathcal{O} \rangle_\theta, \Psi) = \Psi^{eq}(\theta) - \Psi;
\]

where \(\langle \mathcal{O} \rangle_\theta = \{ \langle \mathcal{O}_1 \rangle_\theta, \ldots, \langle \mathcal{O}_n \rangle_\theta \} \).

Applying this proposition to example in section 2, one has the contact Hamiltonian \((34)\). Similar to this proposition, one has the following.

Proposition 4.4 (Dual moment dynamical system as a contact Hamiltonian system). The dynamical system in proposition 3.2 can be written as a contact Hamiltonian system.

Proof. Identify \(x, y, z\) and \(\varpi\) in \((31)\) with \(x^a = \Theta^a, \quad y_\eta = \langle \Theta^a \rangle_\eta, \quad \varpi(x) = \Phi^{eq}(\eta), \quad z = H\).

One then sees that \((31)\) is identical to the dynamical system in proposition 3.2. This system is generated by the contact Hamiltonian \(h\) such that

\[
h(\eta, \langle \Theta \rangle_\eta, \Phi) = \Phi^{eq}(\eta) - \Phi;
\]

where \(\langle \Theta \rangle_\eta = \{ \langle \Theta^1 \rangle_\eta, \ldots, \langle \Theta^n \rangle_\eta \}\), and \(\Phi = H\).

Applying this proposition to example in section 2, one has the contact Hamiltonian \((35)\). The following are geometric descriptions of the propositions 3.1 and 3.2.
1. Let $Y$ be a vector field associated with the dynamical system in proposition 3.1.

$$\frac{\partial}{\partial t} + \psi = Y$$

and $\exp(tX_H): C \to C$ the exponential map associated with $X_H$. Then the point in $C$ at time $t' + t$ is obtained by applying $\exp(t'X_H)$ to $q(t)$. That is, $q(t' + t) = \exp(t'X_H)q(t)$. A set of schematic pictures of this vector field on $C$ is depicted in figure 3.

2. Interchange the dynamical variables with static variables in such a way that

$$\theta \mapsto \Theta(t), \quad \{\theta^a\}_\theta(t) \mapsto \eta^a, \quad a \in \{1, \ldots, n\}.$$ 

Also, $\Psi(\theta, t) \mapsto H(\eta, t)$. Then, let $X_H$ be a vector field associated with the dynamical system in proposition 3.2,

$$\frac{d}{dt} + \frac{\partial}{\partial \eta} = H$$

Let $M$ be a manifold, $\gamma: \mathbb{R} \to M$, $t \mapsto \gamma(t)$ a curve on $M$, $\gamma$ a vector field on $M$ that is obtained as the push-forward $\gamma_*(\partial/\partial t)$, and $g$ a Riemannian metric tensor field. Then the length of $\gamma$ for $t_0 \leq t \leq t_1$ is defined by

$$l(\gamma) = \int_{t_0}^{t_1} \sqrt{g(\gamma(\cdot), \dot{\gamma}(\cdot))} \, dt.$$
With this introduced metric tensor field, one can calculate the length between a state and the target state, where a contact Hamiltonian vector field is identified with \( \dot{\varphi} \) in (42). This can be used to estimate how far states are away from the equilibrium state or a target state.

In [40] a metric tensor field on contact manifolds has been studied. Let \((C, \lambda)\) be a \((2n + 1)\)-dimensional contact manifold, \((x, y, z)\) coordinates such that \( \lambda = dx - y_d dx^a \) with \( x = \{x^1, ..., x^n\} \) and \( y = \{y_1, ..., y_n\} \). One defines

\[
G = \frac{1}{2} [dx^a \otimes dy^b + dy^b \otimes dx^a] + \lambda \otimes \lambda. \tag{43}
\]

This metric tensor field is consistent with the one at the equilibrium state, in the sense that the Fisher metric tensor field on a Legendre submanifold generated by a function is obtained with the pull-back of \( G \).

The following gives the length along the introduced contact Hamiltonian vector field.

**Lemma 4.1.** The length between a state and the equilibrium state for the primary moment dynamical system calculated with (43) is

\[
l[X_\Psi^f]_{\infty} = | \Psi(\theta(t)) - \Psi^e | = | h(\theta, \{ \mathcal{O} \}_\theta, \Psi) |, \tag{44}
\]

where this \( h \) is specified as (38). Then the convergence rate for (44) is exponential.

**Proof.** Since a state is described by \( X_\theta \) in (40). Substituting \( X_\theta \) into (42), one has

\[
l[X_\Psi^f]_{\infty} = \int_{-\infty}^{\infty} \sqrt{G(X_\theta, X_\theta)} \, dt' = \int_{-\infty}^{\infty} \sqrt{\left( \frac{d\Psi}{dt} \right)^{\dot{\varphi}}} \, dt' = | \Psi(t) - \Psi^e |.
\]

With (38), the last equality above can be written in terms of the contact Hamiltonian \( h \) as (44). It follows from (32) that the convergence rate is exponential. \( \square \)

Similarly one has the following.

**Lemma 4.2.** The length between a state and the equilibrium state for the dual moment dynamical system calculated with (43) is

\[
l[X_\Phi^f]_{\infty} = | \Phi(\phi(t)) - \Phi^e | = | h(\phi, \{ \mathcal{O} \}_\phi, \Phi) |, \tag{45}
\]

where this \( h \) is specified as (39). Then the convergence rate for (45) is exponential.

**Proof.** A way to prove this is analogous to the proof of lemma 4.1. \( \square \)

Combining lemmas 4.1, 4.2, and discussions in the previous sections, one has the main theorem in this paper.

**Theorem 4.1** (Contact geometric description of the expectation variables and its convergence). Dynamical systems derived from solvable master equations are described on a contact manifold, and its convergence rate of the associated with the metric tensor field (43) is exponential.

5. Conclusions

This paper offers a viewpoint that sets of dynamical systems exactly derived from master equations can be described in terms of contact geometry and information geometry. To show this explicitly, a solvable toy model as a set of master equations has been concentrated. From this set of master equations and the viewpoint of duality in the sense of information geometry, the dual master equations have been defined. To give a contact geometric description two contact Hamiltonian vector fields have been introduced. By reviewing an existing study, how to describe attractors of such contact Hamiltonian vector fields has been discussed in terms of information geometry. Then, with an introduced metric tensor field, the convergence rate has been shown to be exponential.

Throughout this paper explicit expressions of geometrical objects have been shown by analyzing a spin model.

There are some potential future works that follow from this study.

1. One is to apply the present approach to various Monte Carlo methods. Since the present solvable master equations, the primary and dual master equations, do not show any state transition, associated MCMC methods belong to a special class. For completeness, some other master equations are of interest to explore how much the proposing geometric approach is effective.

2. In addition, in information geometry, some divergences are useful tools to study various objects. For this reason, as a future work, geometry related to divergences will be concentrated.

3. Also, other kinetic models are expected to give physical applications and implications.

4. It is also interesting to introduce more geometric objects, such as connections and other metric tensor fields, to contact manifolds for analyzing nonequilibrium systems.

5. Besides, there have been various works on geometric theories of MCMC methods. Thus it is interesting to see if there is some link between this study and such existing works.

We believe that the elucidation of these remaining questions will develop geometric theory of master equations, that of MCMC methods, and nonequilibrium statistical mechanics.

Acknowledgments

The author S G is partially supported by JPSJ (KAKENHI) grant number 19K03635. The author H H would like to thank Professor Koji Hukushima for useful discussions, also he is partially supported by JSPS (KAKENHI) grant number 17H01793. Both of the authors are partially supported by JST CREST JPMJCR1761.
Appendix. Component expression of contact Hamiltonian vector field

In this appendix, the component expression of contact Hamiltonian vector field is given. Let \( (C, \lambda) \) be an \((2n + 1)\)-dimensional contact manifold with \( \lambda \) being a contact one-form, \((x, y, z)\) a set of coordinates so that \( \lambda = dz - \gamma_i dx^i \) with \( x = \{x^1, \ldots, x^n\}, y = \{y_1, \ldots, y_n\}, t \) a time, and \( h: C \rightarrow \mathbb{R} \) a function called a contact Hamiltonian. Then the local component expression of the contact Hamiltonian vector field is

\[
\frac{d}{dt}x^a = -\frac{\partial h}{\partial y_a}, \quad \frac{d}{dt}y_a = \frac{\partial h}{\partial x^a} + y_i \frac{\partial h}{\partial z}\frac{\partial z}{\partial y_a},
\]

and

\[
\frac{d}{dt}z = h - \gamma_a \frac{\partial h}{\partial y_a}.
\]  

(46)

Note that there are some conventions, and then signs for this expression are not the same as another expression. The set of equations (31) can be derived by substituting (30) into (46). A formulation with contact Hamiltonian systems is a candidate for analytical mechanics of dissipative systems \([41]\). Also, some contact geometric description of information geometry was investigated in \([11]\).

ORCID iDs

Shin-itiro Goto @ https://orcid.org/0000-0002-5249-1054  
Hideitsu Hino @ https://orcid.org/0000-0002-6405-4361

References

[1] Amari S and Nagaoka H 2000 Methods of Information Geometry (Providence, RI, Oxford: American Mathematical Society, Oxford University Press)
[2] Ay N et al 2017 Information geometry (Berlin: Springer)
[3] Hayashi M 2006 Quantum information An introduction (Berlin: Springer)
[4] Naudts J 2011 Generalized Thermodynamics (Berlin: Springer)
[5] Scarfone A M, Matsuzoe H and Wada T 2018 Information geometry of \( k \)-exponential families: dually exponential families
[6] Sando K et al 2019 Information geometry of modulated linear regression Inf. Geom. 2 43–75
[7] Ma W X 2018 Conservation laws by symmetries and adjoint symmetries Discrete Cont. Dyn. Syst. 11 707–21
[8] Arnold V I 1997 Mathematical Methods of Classical Mechanics
[9] da Silva A C 2008 Lectures on Symplectic Geometry 2nd edn (Berlin: Springer)
[10] Goto S 2015 Legendre submanifolds in contact manifolds as attractors and geometric nonequilibrium thermodynamics J. Math. Phys. 56 073301
[11] Goto S 2016 Contact geometric descriptions of vector fields on dually flat spaces and their applications in electric circuit models and nonequilibrium thermodynamics J. Math. Phys. 57 102702
[12] Goto S and Hino H 2019 Expectation variables on a para-contact metric manifold exactly derived from master equations Springer Lecture Notes in Computer Science vol 11712 (Cham: Springer) 239–47 to appear
[13] Mrugula R 1978 Geometrical formulation of equilibrium phenomenological thermodynamics Rep. Math. Phys. 14 419–27
[14] Mrugula R 2000 On contact and metric structures on thermodynamic spaces Saken Koyukuro 1142 167–81
[15] Bravetti A, Lopez-Monsalve C S and Nettel F 2015 Contact symmetries and Hamiltonian thermodynamics Ann. Phys. 361 377–400
[16] Grmela M 2014 Contact geometry of mesoscopic thermodynamics and dynamics Entropy 16 1652–86
[17] Rajeek S G and Hamilton-Jacobi A 2008 formalism for thermodynamics Ann. Phys. 323 2265–85
[18] Bravetti A 2018 Contact geometry and thermodynamics Int. Geom. Methods Mod. Phys. 16 1940003
[19] Glauber R J 1967 Time dependent Statistics of the Ising model J. Math. Phys. 4 294–307
[20] Van Kampen N G 2007 Stochastic Processes in Physics and Chemistry 3rd edn (Amsterdam: North-Holland)
[21] Gardner C 2009 Stochastic Methods 4th edn (Berlin: Springer)
[22] Stathopoulos V and Girolami M A 2013 Markov chain Monte Carlo interference for Markov jump processes via the linear noise approximation Phil. Trans. R. Soc. A 371 20110541
[23] Metropolis N et al 1953 Equation of state calculations by fast computing machines J. Chem. Phys. 21 1087–92
[24] Kirkpatrick S, Gelatt C D and Vecchi M P 1983 Optimization by simulated annealing Science 220 671–80
[25] Iba Y 2001 Extended ensemble Monte Carlo Int. J. Mod. Phys. B 12 1–49
[26] Marinari E and Parisi G 1992 Simulated tempering: a new method and application to spin glass simulations J. Phys. Soc. Japan 65 1604–8
[27] Marinari E and Parisi G 1992 Simulated tempering: a new method and application to spin glass simulations J. Phys. Soc. Japan 65 1604–8
[28] Takabatake K 2004 Information geometry of Gibbs sampler Eurphys. Lett. 19 451–8
[29] Lyubartsev A P et al 1992 New approach to Monte Carlo calculations of the free energy: method of expanded ensembles J. Chem. Phys. 96 1776–83
[30] Hukushima K and Nemoto K 1996 Exchange Monte Carlo method and application to spin glass simulations J. Phys. Soc. Japan 65 1604–8
[31] Fujisawa A and Amari S 1995 Gradient systems in view of information geometry Physica D 80 317–27
[32] Nakamura Y 1994 Gradient system associated with probability distributions Japan. J. Ind. Appl. Math. 11 21–30
[33] Uwano Y 2016 All the trajectories of an extended averaged Hebbian learning equation on the quantum state space are the \( \epsilon \)-geodesics Math. Model. Geom. 4 19–33
[34] Noda T 2011 Symplectic structures on statistical manifolds J. Aust. Math. Soc. 90 371–84
[35] Takabatake K 2004 Information geometry of Gibbs sampler Proc. WSEAS Int. Conf.
[36] Nagaoka H 2005 The exponential family of Markov chains and its information geometry SITA 2005 arXiv:1701.06119
[37] Nielsen F and Hadijere G 2018 Monte Carlo information geometry: the dually flat case arXiv:1803.07225
[38] Mrugula R 1990 Statistical approach to the geometric structure of thermodynamics Phys. Rev. A 41 3156–60
[39] Nakahara M 1990 Geometry, Topology and Physics (Bristol: Institute of Physics Publishing)
[40] Matsuzoe H and Henni M 2013 Hessian structures on deformed exponential families Geometric Science of Information (Lecture Notes in Computer Science) vol 8085 (Berlin: Springer) pp 275–82
[41] Bravetti A 2017 Contact Hamiltonian dynamics: the concept and its use Entropy 19 535
[42] Bravetti A and Lopez-Monsalve C S 2015 Para-Hessian geometry in thermodynamic fluctuation theory J. Phys. A: Math. Theor. 48 125206
[43] Bravetti A, Cruz H and Tapias D 2017 Contact Hamiltonian mechanics Ann. Phys. 376 17–39