

# Unifying various thermodynamically compatible modeling approaches

Liangrong Peng<sup>1,2</sup> and Liu Hong<sup>3,\*</sup>

<sup>1</sup>School of Computer and Data Science, Minjiang University, Fuzhou 350108, China

<sup>2</sup>Minjiang Collaborative Center for Theoretical Physics, Minjiang University, Fuzhou 350108, China

<sup>3</sup>School of Mathematics, Sun Yat-sen University, Guangzhou 510275, China

E-mail: [hongliu@sysu.edu.cn](mailto:hongliu@sysu.edu.cn)

Received 19 August 2025, revised 27 September 2025

Accepted for publication 28 October 2025

Published 26 November 2025



CrossMark

## Abstract

In this study, we propose that many different thermodynamic modeling approaches, including the general equation for the non-equilibrium reversible–irreversible coupling (GENERIC), Onsager’s variational principle, the energetic variational approach, and classical irreversible thermodynamics, can all be cast into the gradient-conservative structure (GCS). GCS enjoys many nice mathematical properties, has close connection with the large deviations principle and gradient flows in Wasserstein space, and fulfills laws of thermodynamics. Our results demonstrate that the GCS may serve as a unified theoretical framework to model various non-equilibrium thermodynamic processes.

Keywords: non-equilibrium thermodynamics, gradient-conservative structure, generalized gradient systems, variational approach

(Some figures may appear in colour only in the online journal)

## 1. Introduction

It is widely known that equilibrium thermodynamics is a well-established field with three laws of thermodynamics as hallmarks. However, most natural systems are not in the equilibrium state. They can exchange energy and matter with the environment and have finite driving forces [1]. To deal with them, we need to extend the concept of equilibrium thermodynamics to non-equilibrium. It is a great pity that the current field of non-equilibrium (or irreversible) thermodynamics has been divided into many different ‘schools’, including classical irreversible thermodynamics [2], rational thermodynamics [3], extended irreversible thermodynamics [4], the general equation for the non-equilibrium reversible-irreversible coupling (GENERIC) [5] and many others. Each school has its own starting point, assumptions, and ways of formulation, but all keep the same aim in mind—to construct thermodynamically compatible and mathematically rigorous models for new physical systems under non-equilibrium situations.

The failure of previous studies in constructing a unified general framework for non-equilibrium thermodynamics stems from the following challenges:

- Non-equilibrium phenomena, like bistability, memory-effect, and pattern formation, are far more complicated than those under the equilibrium state;
- Non-equilibrium systems are time-evolving and do not meet the requirement of time reversibility, meaning modeling by dynamical equations are more suitable (or easier) than a thermodynamic formulation;
- Many non-equilibrium phenomena are system-size dependent and cannot be dealt with from a macroscopic thermodynamic view;
- A microscopic theory for non-equilibrium systems, in analogy to Boltzmann’s statistical mechanics for equilibrium systems, is still lacking.

Motivated by the gradient-conservative structure proposed by Mielke *et al* [6], in this study we propose that many different thermodynamic modeling approaches, including GENERIC [5], Onsager’s variational principle [7], the energetic variational approach [8], and classical irreversible thermodynamics [4], can all be cast into the gradient-conservative structure, which therefore suggests a unified theoretical framework to model various non-equilibrium thermodynamic processes.

To be exact, in section 2 we first revisit the basic formulation of the gradient-conservative structure, its nice

\* Author to whom any correspondence should be addressed.

mathematical properties, close relations to the large deviations principle, and interpretation in Wasserstein space. Our main result, the unification of various thermodynamically compatible modeling approaches used in the field of non-equilibrium thermodynamics in the sense of the gradient-conservative structure, is demonstrated in section 3. Section 4 includes two concrete examples, through which the consistency of different modeling approaches will be fully appreciated. The last section contains a brief conclusion and some preliminary discussion.

## 2. Gradient-conservative structure

### 2.1. Basic formulation

We introduce vector  $\mathbf{z} = \mathbf{z}(t)$  as a group of state variables on a Riemann manifold  $\mathcal{Z}$ , whose time evolution is assumed to obey the following gradient-conservative structure (GCS) [9],

$$\frac{d\mathbf{z}}{dt} = \nabla_{\xi} \psi^*(\mathbf{z}, \xi)|_{\xi=-\nabla F(\mathbf{z})} + \mathbf{A}(\mathbf{z}) \cdot \nabla E(\mathbf{z}). \quad (1)$$

Here, the first term on the right-hand side  $\nabla_{\xi} \psi^*(\mathbf{z}, \xi)|_{\xi=-\nabla F(\mathbf{z})}$  denotes the gradient dynamics with a non-negative dissipation function  $\psi^*(\mathbf{z}, \xi)$ , whose minimum is attained at  $\xi = \mathbf{0}$  ( $\psi^*(\mathbf{z}, \mathbf{0}) = 0$ ), while the second term  $\mathbf{A}(\mathbf{z}) \cdot \nabla E(\mathbf{z})$  presents the conservative dynamics.  $\mathbf{A}^{\mathbb{T}} = -\mathbf{A}$  is an anti-symmetric operator.  $E(\mathbf{z})$  is the total energy of the system,  $F(\mathbf{z})$  is the relative entropy or free energy function, where  $\nabla E$  and  $\nabla F$  denote the gradients of energy and free energy with respect to state variables  $\mathbf{z}$ , respectively.

Along with the gradient-conservative structure, two transverse conditions are further required, i.e.

$$\nabla E(\mathbf{z}) \cdot \frac{\partial \psi^*(\mathbf{z}, \xi)}{\partial \xi} \Big|_{\xi=-\nabla F} = 0, \quad (2)$$

$$\nabla F(\mathbf{z}) \cdot \mathbf{A}(\mathbf{z}) \cdot \nabla E(\mathbf{z}) = 0. \quad (3)$$

As we will see, these two conditions guarantee that the system dynamics is compatible with the general laws of thermodynamics, that is, the energy conservation and relative entropy decreasing in time.

$$\frac{dE(\mathbf{z})}{dt} = \nabla E(\mathbf{z}) \cdot \frac{d\mathbf{z}}{dt} = \nabla E(\mathbf{z}) \cdot \mathbf{A}(\mathbf{z}) \cdot \nabla E(\mathbf{z}) = 0, \quad (4)$$

$$\begin{aligned} \frac{dF(\mathbf{z})}{dt} &= \nabla F(\mathbf{z}) \cdot \frac{\partial \psi^*(\mathbf{z}, \xi)}{\partial \xi} \Big|_{\xi=-\nabla F} \\ &= -[\psi^*(\mathbf{z}, -\nabla F) + \psi(\mathbf{z}, \dot{\mathbf{z}})] \leq 0, \end{aligned} \quad (5)$$

where  $\dot{\mathbf{z}} = \frac{\partial \psi^*(\mathbf{z}, \xi)}{\partial \xi} \Big|_{\xi=-\nabla F}$  denotes the part of the gradient dynamics, and  $\psi(\mathbf{z}, \dot{\mathbf{z}})$  is the Legendre-Fenchel transform of  $\psi^*(\mathbf{z}, -\nabla F)$ . The second equality in equation (5) holds due to the Legendre-Fenchel transform, and the last inequality holds due to the non-negativity of  $\psi^*$  and  $\psi$ , which will be further clarified below equation (7).

It is important to note that the GCS, by construction, describes systems whose dynamics are thermodynamically compatible. The transverse conditions (2) and (3) are necessary and sufficient to enforce energy conservation and relative entropy decrease. The GCS finds applications in various systems, including the chemical reactions and diffusion equations discussed in section 4. Although GCS provides a general framework, it imposes specific structural constraints related to thermodynamic compatibility that not all dynamical systems satisfy. To illustrate this fact, we present a detailed counter-example.

Consider a damped harmonic oscillator described by  $dq/dt = p$  and  $dp/dt = -q - \gamma p$ , where  $\mathbf{z} = (q, p)^{\mathbb{T}}$  represents the phase space coordinates and  $\gamma > 0$  is the damping coefficient. If we attempt to represent this system within the GCS framework using the natural energy function  $E = \frac{1}{2}(p^2 + q^2)$  and the standard antisymmetric matrix  $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ , the dissipative component must satisfy  $(\nabla_{\xi} \psi^*(\mathbf{z}, \xi))_{\xi=-\nabla F(\mathbf{z})} = d\mathbf{z}/dt - \mathbf{A}(\mathbf{z}) \cdot \nabla E(\mathbf{z}) = (0, -\gamma p)^{\mathbb{T}}$ . This representation violates the transverse condition  $\nabla E(\mathbf{z}) \cdot (\nabla_{\xi} \psi^*(\mathbf{z}, \xi))_{\xi=-\nabla F(\mathbf{z})} = 0$ , since the dot product yields  $-\gamma p^2 \neq 0$  for non-zero  $p$ . Alternatively, if we set  $\mathbf{A} = \mathbf{0}$ , the system would need to be a pure gradient flow. However, the vector field  $(p, -q - \gamma p)^{\mathbb{T}}$  is not conservative (its curl is non-zero), precluding this representation. Thus, the damped harmonic oscillator cannot be simply cast into the standard GCS framework, as illustrated by the two methods proposed above.

Overall, the conservative part in the GCS defines a dynamics that conserves energy as equation (4). If the anti-symmetric operator  $\mathbf{A}(\mathbf{z})$  further satisfies the Jacobi identity, this part defines a Hamiltonian structure, which is familiar to physicists and mathematicians and will not be discussed further here. In contrast, there have been numerous significant mathematical advances concerning the dissipative part given through a gradient structure in recent years, which therefore is worthy of further discussion. To be specific, we will show how to extend the gradient structure into a much generalized system, and its intrinsic relation with stochastic processes through the large deviations principle, as well as an elegant interpretation in Wasserstein space instead of the classical Euclidean space.

### 2.2. Generalized gradient systems

The gradient structure in equation (1) can be introduced in a much broader way. The triple  $(\mathbf{z}, \psi^*, F)$  is called a generalized gradient system [6, 10], if it satisfies

$$\psi(\mathbf{z}, \dot{\mathbf{z}}) + \psi^*(\mathbf{z}, -\nabla F) + \langle \nabla F, \dot{\mathbf{z}} \rangle = 0, \quad (6)$$

where the relative entropy  $F: \mathcal{Z} \rightarrow \mathbb{R}$  for  $\forall \mathbf{z} \in \mathcal{Z}$ ,  $\langle \cdot, \cdot \rangle$  represents the inner product. We denote the tangent and cotangent spaces of the Riemann manifold  $\mathcal{Z}$  as  $T\mathcal{Z}$  and  $T^*\mathcal{Z}$  separately.  $\psi^*$  and  $\psi$  constitute a Legendre-Fenchel duality

pair,

$$\begin{cases} \psi(\mathbf{z}, \zeta) = \sup_{\xi \in T_z^* \mathcal{Z}} (\langle \xi, \zeta \rangle - \psi^*(\mathbf{z}, \xi)), \\ \psi^*(\mathbf{z}, \xi) = \sup_{\zeta \in T_z \mathcal{Z}} (\langle \xi, \zeta \rangle - \psi(\mathbf{z}, \zeta)). \end{cases} \quad (7)$$

If  $\psi^*$  is strictly convex and differentiable in  $\xi$ , the supremum in the first equation of (7) is attained at a unique  $\bar{\xi} \in T_z^* \mathcal{Z}$  satisfying the stationary condition  $\zeta = \left( \frac{\partial \psi^*(\mathbf{z}, \xi)}{\partial \xi} \right)_{\xi=\bar{\xi}(\mathbf{z}, \zeta)}$ , which implies the identity:

$$\psi(\mathbf{z}, \zeta) = \langle \bar{\xi}(\mathbf{z}, \zeta), \zeta \rangle - \psi^*(\mathbf{z}, \bar{\xi}(\mathbf{z}, \zeta)).$$

As defined via a Legendre–Fenchel transform,  $\psi$  is guaranteed to be a convex function of  $\zeta$ . A similar conclusion holds for the second equation of (7) too. According to the Fenchel–Young inequality, we always have

$$\psi(\mathbf{z}, \dot{\mathbf{z}}) + \psi^*(\mathbf{z}, -\nabla F) + \langle \nabla F, \dot{\mathbf{z}} \rangle \geq 0, \quad (8)$$

meaning that equation (6) actually implies an equivalent form

$$\begin{cases} \dot{\mathbf{z}}(t) = \frac{\partial \psi^*(\mathbf{z}, \xi)}{\partial \xi} |_{\xi=-\nabla F}, \\ \nabla F(\mathbf{z}) = -\frac{\partial \psi(\mathbf{z}, \zeta)}{\partial \zeta} |_{\zeta=\dot{\mathbf{z}}}. \end{cases} \quad (9)$$

It is recognized that the first equation gives the dissipative part in the GCS.

A generalized gradient system further requires that

1.  $\psi^*(\mathbf{z}, \xi)$  is convex with respect to the second argument  $\xi$ ;
2.  $\psi^*(\mathbf{z}, \xi) \geq 0$  for  $\forall \xi \in T_z^* \mathcal{Z}$ , and  $\psi^*(\mathbf{z}, \mathbf{0}) = \mathbf{0}$ ;
3.  $\psi^*(\mathbf{z}, \xi) = \psi^*(\mathbf{z}, -\xi)$  for  $\forall \mathbf{z} \in \mathcal{Z}, \forall \xi \in T_z^* \mathcal{Z}$ .

The second requirement on  $\psi^*$  will apply a dual requirement on  $\psi$ , since

$$\begin{cases} \psi^*(\mathbf{z}, \mathbf{0}) = 0 \Leftrightarrow \psi(\mathbf{z}, \zeta) \geq 0, & \forall \zeta \in T_z \mathcal{Z}, \\ \psi(\mathbf{z}, \mathbf{0}) = 0 \Leftrightarrow \psi^*(\mathbf{z}, \xi) \geq 0, & \forall \xi \in T_z^* \mathcal{Z}. \end{cases}$$

Now it is direct to show that

$$\frac{dF(\mathbf{z})}{dt} = \langle \nabla F, \dot{\mathbf{z}} \rangle = -[\psi(\mathbf{z}, \dot{\mathbf{z}}) + \psi^*(\mathbf{z}, -\nabla F)] \leq 0,$$

a special case of general results on the GCS. The last requirement is equivalent to the time symmetry requirement on  $\psi$ , that is,  $\psi(\mathbf{z}, \zeta) = \psi(\mathbf{z}, -\zeta)$  for all  $\zeta \in T_z \mathcal{Z}$ . It is natural in many systems, in which velocities with opposite directions lead to the same dissipation.

### 2.3. Relation to the large deviations principle

The generalized gradient system (6) has a close connection with the underlying stochastic processes, which is clarified through the principle of large deviations. Consider the following dynamics

$$\begin{cases} \frac{d}{dt} \mathbf{z}(t) = \mathbf{R}(\mathbf{z}), \\ \mathbf{z}(0) = \mathbf{z}_0, \end{cases} \quad (10)$$

whose solution is given by  $\{\mathbf{z}(t)\}_{0 \leq t \leq T}$ . Further, suppose

the above dynamics is invariant under the time-reversal transformation.

Let us make small random perturbations of  $\{\mathbf{z}(t)\}_{0 \leq t \leq T}$  and consider the corresponding stochastic process  $\{\mathbf{z}_\epsilon(t)\}_{0 \leq t \leq T}$ , where  $0 < \epsilon \ll 1$  denotes the noise level. Following the philosophy of the Fredlin–Wentzell theory [11], we assume that the large deviations principle holds for  $\{\mathbf{z}_\epsilon(t)\}_{0 \leq t \leq T}$  under the condition of fixed  $\mathbf{z}_\epsilon(0) = \mathbf{z}_0$  and  $\mathbf{z}_\epsilon(T) = \mathbf{z}_T$  and as  $\epsilon \rightarrow 0$

$$\mathbb{P}(\{\mathbf{z}_\epsilon(t)\}_{0 \leq t \leq T} = \{\mathbf{z}(t)\}_{0 \leq t \leq T}) \asymp \exp \left[ -\frac{1}{\epsilon} \int_0^T L(\mathbf{z}, \dot{\mathbf{z}}) dt \right], \quad (11)$$

in which the time integral of the Lagrangian gives the large deviations rate function (or action in physics). Notice that, in the limit of  $\epsilon \rightarrow 0$ , the stochastic differential equation (SDE)

$$d\mathbf{z}_t = \mathbf{R}(\mathbf{z}_t) dt + \sqrt{2\epsilon} d\mathbf{W}_t,$$

is not the ‘straightforward’ first-order ordinary differential equation (ODE)  $d\mathbf{z}/dt = \mathbf{R}(\mathbf{z})$  but instead a second-order ODE from the Euler–Lagrange equation associated with  $L(\mathbf{z}, \dot{\mathbf{z}})$  with boundary conditions at  $\mathbf{z}(0)$  and  $\mathbf{z}(T)$ .

Furthermore, we assume that the Lagrangian is time reversible, meaning there exists a function  $U = U(\mathbf{z})$  such that [6, 12]

$$L(\mathbf{z}, \dot{\mathbf{z}}) - L(\mathbf{z}, -\dot{\mathbf{z}}) = 2\langle \nabla U, \dot{\mathbf{z}} \rangle, \quad (12)$$

$\forall \mathbf{z} \in \mathcal{Z}, \forall \dot{\mathbf{z}} \in T_z \mathcal{Z}$ . Now we are ready to introduce the dissipation function:

$$\psi^*(\mathbf{z}, \xi) = H(\mathbf{z}, \nabla U + \xi) - H(\mathbf{z}, \nabla U), \quad (13)$$

in which the Hamiltonian function is defined as the Legendre–Fenchel transformation of the Lagrangian function,

$$H(\mathbf{z}, \xi) = \sup_{\dot{\mathbf{z}} \in T_z \mathcal{Z}} (\langle \xi, \dot{\mathbf{z}} \rangle - L(\mathbf{z}, \dot{\mathbf{z}})). \quad (14)$$

Most importantly, the dissipation function defined above satisfies

1.  $\psi^*(\mathbf{z}, \xi)$  is convex with respect to the second argument  $\xi$ ;
2.  $\psi^*(\mathbf{z}, \xi) \geq 0 \forall \xi \in T_z^* \mathcal{Z}$ , and  $\psi^*(\mathbf{z}, \mathbf{0}) = 0$ ;
3.  $\psi^*(\mathbf{z}, \xi) = \psi^*(\mathbf{z}, -\xi) \forall \xi \in T_z^* \mathcal{Z}$ ,

by recalling the mathematical properties of the large deviations rate function as well as the Legendre–Fenchel transformation. See Appendix A for a proof.

Defining  $\psi(\mathbf{z}, \dot{\mathbf{z}}) = \sup_{\xi \in T_z^* \mathcal{Z}} (\langle \xi, \dot{\mathbf{z}} \rangle - \psi^*(\mathbf{z}, \xi))$ , it is straightforward to verify that the Lagrangian function has an alternative insightful expression through the dissipation function, i.e.

$$L(\mathbf{z}, \dot{\mathbf{z}}) = \psi(\mathbf{z}, \dot{\mathbf{z}}) + \psi^*(\mathbf{z}, -\nabla U) + \langle \nabla U, \dot{\mathbf{z}} \rangle \geq 0, \quad (15)$$

which is non-negative due to the Fenchel–Young inequality. In particular, when  $L(\mathbf{z}, \dot{\mathbf{z}}) = 0$ , we have

$$\psi(\mathbf{z}, \dot{\mathbf{z}}) + \psi^*(\mathbf{z}, -\nabla U) + \langle \nabla U, \dot{\mathbf{z}} \rangle = 0,$$

meaning  $(\mathbf{z}, \psi^*, U)$  defines a generalized gradient system. As the large deviations rate function equals zero, the stochastic process  $\{\mathbf{z}_\epsilon(t)\}_{0 \leq t \leq T}$  converges to the deterministic one with

probability one. It can be rewritten into the form of gradient flows, i.e.

$$\dot{z}(t) = \frac{\partial \psi^*(z, \xi)}{\partial \xi} \Big|_{\xi = -\nabla U}.$$

Therefore, the Lagrangian function  $L(z, \dot{z}) \geq 0$  in equation (15) provides a natural measurement of the deviations of stochastic processes from the deterministic gradient flow.

### 2.4. Interpretation of gradient flows in Wasserstein space

The universality of gradient structures was fully appreciated through the milestone work of Otto [13], in which he, for the first time, pointed out that many non-gradient dynamics could be reformulated into a gradient one by making a change in the metric of the space. To see this point, we turn to an alternative interpretation of gradient flows in Wasserstein space.

**Definition 1.** Given two functions  $h_1, h_2: \Omega \rightarrow \mathbb{R}$  with  $\int_{\Omega} h_1 dx = \int_{\Omega} h_2 dx = 0$ , one can define their Wasserstein scalar product at  $\rho$  as

$$\langle h_1, h_2 \rangle_{\rho} := \int_{\Omega} (\nabla_x \psi_1 \cdot \nabla_x \psi_2) \rho dx,$$

where

$$\begin{cases} \nabla_x \cdot (\rho \nabla_x \psi_i) = -h_i \text{ in } \Omega \\ \vec{n} \cdot \nabla_x \psi_i = 0 \text{ on } \partial\Omega. \end{cases}$$

$\nabla_x$  denotes the normal gradient in Euclidean space, and  $\vec{n}$  denotes the boundary normal vector.

Now we can define the gradient of a functional in Wasserstein space.

**Definition 2.** Given a functional  $F: \mathcal{P}_2(\Omega) \rightarrow \mathbb{R}$ , its gradient with respect to the Wasserstein scalar product at  $\bar{\rho} \in \mathcal{P}_2(\Omega)$  is the unique function  $\nabla_{W_2} F(\bar{\rho})$  (if it exists) such that

$$\left\langle \nabla_{W_2} F(\bar{\rho}), \frac{\partial \rho_{\epsilon}}{\partial \epsilon} \Big|_{\epsilon=0} \Big|_{\bar{\rho}} \right\rangle = \frac{d}{d\epsilon} \Big|_{\epsilon=0} F(\rho_{\epsilon}),$$

for any smooth curve  $\rho_{\epsilon}: (-\epsilon_0, \epsilon_0) \rightarrow \mathcal{P}(\Omega)$  with  $\rho_0 = \bar{\rho}$ .

Now with respect to functional  $F: \mathcal{P}_2(\Omega) \rightarrow \mathbb{R}$  and the probability measure  $\bar{\rho} \in \mathcal{P}_2(\Omega)$ , let us denote by  $\delta F(\bar{\rho})/\delta \rho$  its first  $\mathcal{L}^2$ -variation, which is the function in  $\mathcal{L}^2(\Omega)$  such that

$$\frac{d}{d\epsilon} \Big|_{\epsilon=0} F(\rho_{\epsilon}) = \int_{\Omega} \frac{\delta F(\bar{\rho})}{\delta \rho}(\mathbf{x}) \frac{\partial \rho_{\epsilon}(\mathbf{x})}{\partial \epsilon} \Big|_{\epsilon=0} dx. \tag{16}$$

Therefore, by the definition of the Wasserstein scalar product, we can deduce that

$$\nabla_{W_2} F(\bar{\rho}) = -\nabla_x \cdot \left( \nabla_x \left( \frac{\delta F(\bar{\rho})}{\delta \rho} \right) \bar{\rho} \right). \tag{17}$$

Now we are ready to introduce gradient flows in Wasserstein space.

**Definition 3.** Given a functional  $F: \mathcal{P}_2(\Omega) \rightarrow \mathbb{R}$ , a curve of probability measure  $\rho: [0, T) \rightarrow \mathcal{P}_2(\Omega)$  is a gradient flow of  $F$  with respect to  $W_2$  and with starting point  $\bar{\rho}_0$  if

$$\begin{cases} \frac{\partial \rho_t}{\partial t} = -\nabla_{W_2} F(\rho_t), \\ \rho_0 = \bar{\rho}_0, \end{cases}$$

Considering the above definition of gradient flows and concrete formulas for the Wasserstein gradient of  $F$ , we could rewrite plenty of classical examples in mathematical physics into gradient flows in Wasserstein space [14].

- i) Given  $F(\rho) = \int \rho \ln \rho dx$ , the Wasserstein gradient flow leads to the heat equation  $\frac{\partial \rho}{\partial t} = -\nabla_{W_2} \mathcal{F}[\rho] = \Delta_x \rho$ .
- ii) Given  $F(\rho) = \int [\rho \ln \rho + \rho V(\mathbf{x})] dx$ , the gradient flow gives the Fokker–Planck equation  $\frac{\partial \rho}{\partial t} = \Delta_x \rho + \nabla_x \cdot (\rho \nabla_x V)$ .
- iii) Given  $F(\rho) = \frac{1}{m-1} \int \rho^m dx$  for  $m > 1$ , the gradient flow gives the porous medium equation  $\frac{\partial \rho}{\partial t} = \Delta_x(\rho^m)$ .
- iv) Given  $F(\rho) = \int [U(\rho) + \rho V(\mathbf{x})] dx + \frac{1}{2} \int \int W(\mathbf{x} - \mathbf{y}) \rho(\mathbf{x})(\mathbf{y}) dx dy$ , the gradient flow gives the diffusion–advection–aggregation model  $\frac{\partial \rho}{\partial t} = \nabla_x \cdot \{ \rho [\nabla_x U'(\rho) + \nabla_x V(\mathbf{x}) + (\nabla_x W) * \rho] \}$ , where  $U'(\rho) = \frac{\delta U(\rho)}{\delta \rho}$ ,  $W * \rho = \int W(\mathbf{x} - \mathbf{y}) \rho(\mathbf{y}) dy$ .

## 3. Unifying various thermodynamically compatible modeling approaches

### 3.1. GENERIC

The GENERIC, short for ‘general equation for non-equilibrium reversible–irreversible coupling’, was first proposed by Grmela in 1989 [15], then developed by Grmela and Öttinger in 1997 [5, 16] for describing the time evolution of physical systems out of equilibrium.

The fundamental equation for the GENERIC looks quite similar to that of the GCS [17],

$$\frac{dz}{dt} = \mathbf{M}(z) \cdot \nabla S(z) + \mathbf{A}(z) \cdot \nabla E(z). \tag{18}$$

The two contributions to the time evolution of  $z$  generated by the entropy  $S$  and the total energy  $E$  above are the irreversible and reversible contributions to dynamics, respectively. The so-called Poisson operator  $\mathbf{A}$  is antisymmetric, whereas the friction operator  $\mathbf{M}$  is both symmetric and positive semi-definite. Furthermore, the entropy function  $S(z)$  is assumed to a convex function.

Equation (18) is supplemented by two complementary degeneracy requirements,

$$\mathbf{M}(z) \cdot \nabla E(z) = 0, \tag{19}$$

$$\mathbf{A}(z) \cdot \nabla S(z) = 0. \tag{20}$$

The first requirement that the energy gradient  $\nabla E(z)$  is in the null-space of the friction operator  $\mathbf{M}$  expresses the

conservation of the total energy in a closed system by the irreversible contribution to the dynamics. On the other hand, the second requirement that the entropy gradient  $\nabla S(\mathbf{z})$  is in the null-space of the Poisson operator  $\mathbf{A}$  expresses the reversible nature of the second term in equation (18), independent of the particular form of the Hamiltonian.

Under the help of degeneracy requirements, it becomes straightforward to show  $dE/dt = 0$  and  $dS/dt \geq 0$ , in accordance with the first and second laws of thermodynamics. Therefore, there is a nice one-to-one correspondence between the degeneracy requirements in the GENERIC and the transverse conditions in the GCS.

Provided that the dissipation function takes a quadratic form of  $\xi = -\nabla F$ , i.e.  $\psi^*(\mathbf{z}, \xi) = \frac{1}{2}\langle \xi, \mathbf{M}(\mathbf{z})\xi \rangle$ , and replacing the relative entropy function  $F$  by its counterpart, the negative entropy function,  $-S$ , we see the GCS turns into the GENERIC mentioned above. Actually, the gradient structure has also been widely used to reformulate the GENERIC in recent years [9].

### 3.2. Onsager's variational principle

Onsager's variational principle (OVP) was originally proposed by Lars Onsager in 1931. In two seminal papers [18, 19], he not only established the famous reciprocal relation but also proposed the OVP as a variational principle which is equivalent to the linear force–flux relations in describing irreversible processes.

For isothermal systems, the OVP assumes the state variables follow the dynamic path that minimizes the so-called Rayleighian by Masao Doi [7], i.e.

$$\mathcal{R}(\mathbf{z}, \dot{\mathbf{z}}) = \psi(\mathbf{z}, \dot{\mathbf{z}}) + \dot{F}(\mathbf{z}, \dot{\mathbf{z}}). \quad (21)$$

Here  $\psi(\mathbf{z}, \dot{\mathbf{z}})$  is the dissipation function, which is usually taken as a quadratic function of  $\dot{\mathbf{z}}$ , i.e.  $\psi(\mathbf{z}, \dot{\mathbf{z}}) = \frac{1}{2}\langle \dot{\mathbf{z}}, \mathbf{M}(\mathbf{z})\dot{\mathbf{z}} \rangle$ .  $\mathbf{M}(\mathbf{z})$  is a positive definite operator and symmetric due to Onsager's reciprocal relations.  $\dot{F}(\mathbf{z}, \dot{\mathbf{z}}) = \langle \nabla F(\mathbf{z}), \dot{\mathbf{z}} \rangle$  represents the change rate of free energy. Then the Euler–Lagrange equation for minimizing  $\mathcal{R}$  with respect to  $\dot{\mathbf{z}}$  is given by [20, 21]

$$\mathbf{M}(\mathbf{z})\dot{\mathbf{z}} = -\nabla F(\mathbf{z}). \quad (22)$$

Therefore, the OVP can be regarded as an extension of the principle of the least dissipation of energy proposed by Lord Rayleigh [22].

The statistical mechanical foundation of OVP was established by Onsager and Machlup in 1953 [23], which is known as the Onsager–Machlup variational principle [21]. The latter states that, with respect to the overdamped Langevin equation, the most probable transition path occurring between state  $\mathbf{z}(t)$  and its nearby state  $\tilde{\mathbf{z}}(t)$  is the one minimizing the Onsager–Machlup function or equivalently the Rayleighian function [24].

The relation between Onsager's variational principle and generalized gradient systems is straightforward. Consider the Lagrangian function

$$L(\mathbf{z}, \dot{\mathbf{z}}) = \psi(\mathbf{z}, \dot{\mathbf{z}}) + \langle \nabla F, \dot{\mathbf{z}} \rangle + \psi^*(\mathbf{z}, -\nabla F) \geq 0,$$

defined in equation (15), which is non-negative due to the Fenchel–Young inequality. Denote the sum of the first two terms as the Rayleighian

$$\mathcal{R}(\mathbf{z}, \dot{\mathbf{z}}) = \psi(\mathbf{z}, \dot{\mathbf{z}}) + \langle \nabla F, \dot{\mathbf{z}} \rangle, \quad (23)$$

in which the second term is exactly the change rate of free energy. It is apparent that the condition for generalized gradient flows ( $L(\mathbf{z}, \dot{\mathbf{z}}) = 0$ ) implies Onsager's variational principle,

$$\frac{\partial L(\mathbf{z}, \dot{\mathbf{z}})}{\partial \dot{\mathbf{z}}} = \frac{\partial \mathcal{R}(\mathbf{z}, \dot{\mathbf{z}})}{\partial \dot{\mathbf{z}}} = 0. \quad (24)$$

### 3.3. Energetic variational approach

The energetic variational approach (EVA) was proposed by Liu *et al* for modeling complex fluids systems in physics, chemistry, and biochemistry [8, 25–27]. The starting point of EVA is an energy-dissipation law

$$\frac{dE}{dt} = -\Delta_E \leq 0, \quad (25)$$

where  $E$  denotes the total energy of the system, and  $\Delta_E$  represents the rate of energy dissipation. It can be obtained by a combination of the first and second laws of thermodynamics for an isothermal and mechanically isolated system, that is

$$\frac{d}{dt}(K + U) = \dot{W} + \dot{Q}, \quad (26)$$

$$T \frac{dS}{dt} = \dot{Q} + \Delta_S. \quad (27)$$

Here  $K$ ,  $U$ ,  $S$ , and  $T$  represent the kinetic energy, internal energy, entropy, and absolute temperature.  $\dot{W}$ ,  $\dot{Q}$ ,  $\Delta_S$  denote the rates of the external work, heat absorption, and entropy production, separately. By subtracting the above two equations and using the isothermal and mechanically isolated conditions  $dT/dt = \dot{W} = 0$ , we can derive the energy-dissipation law, equation (25), with  $\Delta_E = \Delta_S$  and  $E = K + U$ , where  $F = U - TS$  is the Helmholtz free energy.

With respect to the energy-dissipation law in equation (25), the EVA assumes that the dynamics of a system is determined by a combination of the least action principle (LAP) and the maximum dissipation principle (MDP). The LAP, which states the conservative force of a Hamiltonian system, can be derived by taking the variation of the action functional  $A_F = \int_0^t (K - F)d\tau$  with respect to the state variable  $\mathbf{z}$ . Meanwhile, the MDP shows that the dissipation force is given by a variation of the dissipation potential  $\psi$  with respect to the velocity  $\dot{\mathbf{z}}$ . In turn, the force balance condition leads to the underlying evolution equation of the system

$$\frac{\partial A_F(\mathbf{z}, \dot{\mathbf{z}})}{\partial \mathbf{z}} = \frac{\partial \psi(\mathbf{z}, \dot{\mathbf{z}})}{\partial \dot{\mathbf{z}}}. \quad (28)$$

The relation between EVA and the generalized gradient system can be understood through the following discussion. Consider the Lagrangian function

$$L(\mathbf{z}, \dot{\mathbf{z}}) = \psi(\mathbf{z}, \dot{\mathbf{z}}) + \langle \boldsymbol{\xi}, \dot{\mathbf{z}} \rangle + \psi^*(\mathbf{z}, -\boldsymbol{\xi}),$$

whose minimization gives

$$\frac{\partial L(\mathbf{z}, \dot{\mathbf{z}})}{\partial \dot{\mathbf{z}}} = \frac{\partial \psi(\mathbf{z}, \dot{\mathbf{z}})}{\partial \dot{\mathbf{z}}} + \boldsymbol{\xi} = \mathbf{0}. \quad (29)$$

Provided the conservative force  $\boldsymbol{\xi} = -\frac{\partial A_F}{\partial \mathbf{z}}$ , we arrive at the central equation of the EVA. A further comparison between equations (24) and (29) demonstrates the close connection between Onsager's variational approach and the EVA. Both of them take a quite similar form, except that the latter has an addition term,  $-\frac{\partial K}{\partial \mathbf{z}}$ .

### 3.4. Classical irreversible thermodynamics and beyond

Classical irreversible thermodynamics (CIT) also started with the pioneering works of Onsager [18, 19] and later was formulated into a general framework for modeling various irreversible processes by Glandsdorff, Prigogine, de Groot, Mazur, and many others [2]. The theoretical foundation of CIT lays on two key assumptions: (a) the local equilibrium hypothesis, which secures a spatiotemporal entropy density function; (b) a linear thermodynamic force–flux relationship.

To derive thermodynamically compatible equations for the time evolution of state variables  $\mathbf{z}(t)$ , CIT assumes the existence of a convex entropy function  $S(t) = S(\mathbf{z}(t))$  with respect to  $\mathbf{z}(t)$ . Then the famous Gibbs relation gives

$$\frac{dS}{dt} = \frac{1}{T} \left( \frac{dU}{dt} - \frac{dF}{dt} \right) = \frac{1}{T} \left( \frac{\partial U}{\partial \mathbf{z}} \frac{d\mathbf{z}}{dt} - \frac{\partial F}{\partial \mathbf{z}} \frac{d\mathbf{z}}{dt} \right), \quad (30)$$

where  $U$  and  $F$  denote the internal energy and free energy, respectively.  $T$  stands for the temperature, which is a constant.

Using the decomposition

$$\frac{d\mathbf{z}}{dt} = G(\mathbf{z}) + C(\mathbf{z}),$$

where  $G(\mathbf{z})$  and  $C(\mathbf{z})$  represent the gradient dynamics and conservative dynamics respectively, we have

$$\begin{aligned} \frac{dS}{dt} &= \frac{1}{T} \left( \frac{\partial U}{\partial \mathbf{z}} G(\mathbf{z}) + \frac{\partial U}{\partial \mathbf{z}} C(\mathbf{z}) - \frac{\partial F}{\partial \mathbf{z}} C(\mathbf{z}) \right) - \frac{1}{T} \frac{\partial F}{\partial \mathbf{z}} G(\mathbf{z}), \\ &= J_f + \text{epr}, \end{aligned} \quad (31)$$

where  $J_f$  and  $\text{epr}$  represent the entropy exchange rate and entropy production rate separately.

Inspired by the formulation of the GENERIC, we set

$$\mathcal{G}(\mathbf{z}) = -\mathbf{M}(\mathbf{z}) \cdot \frac{\partial F}{\partial \mathbf{z}}, \quad C(\mathbf{z}) = \mathbf{A}(\mathbf{z}) \cdot \frac{\partial U}{\partial \mathbf{z}}, \quad (32)$$

where  $\mathbf{M}(\mathbf{z}) = \mathbf{M}^T(\mathbf{z}) \geq \mathbf{0}$ ,  $\mathbf{A}(\mathbf{z}) = -\mathbf{A}^T(\mathbf{z})$ , and  $\mathbf{M}(\mathbf{z}) \cdot \frac{\partial U}{\partial \mathbf{z}} = \mathbf{A}(\mathbf{z}) \cdot \frac{\partial F}{\partial \mathbf{z}} = \mathbf{0}$ . By making use of formulas above, it is

straightforward to see that

$$\begin{aligned} \frac{\partial U}{\partial \mathbf{z}} \cdot \mathbf{M}(\mathbf{z}) \cdot \frac{\partial F}{\partial \mathbf{z}} &= \frac{\partial F}{\partial \mathbf{z}} \cdot \mathbf{A}(\mathbf{z}) \cdot \frac{\partial U}{\partial \mathbf{z}} \\ &= \frac{\partial U}{\partial \mathbf{z}} \cdot \mathbf{A}(\mathbf{z}) \cdot \frac{\partial U}{\partial \mathbf{z}} = 0. \end{aligned}$$

So, we have

$$J_f = 0, \quad \text{epr} = \frac{1}{T} \left( \frac{\partial F}{\partial \mathbf{z}} \cdot \mathbf{M}(\mathbf{z}) \cdot \frac{\partial F}{\partial \mathbf{z}} \right) \geq 0. \quad (33)$$

It is noted that the non-negativity of the entropy production rate is consistent with the second law of thermodynamics. Therefore, we can conclude that the GCS formulation can be regarded as a special realization of CIT, which is deeply rooted in the principle of large deviations [28].

Extended irreversible thermodynamics (EIT) was proposed by Müller, Jou, Casas-Vázquez, and Lebon as a generalization of CIT, which aims to abandon the local equilibrium hypothesis and deal with high-frequency, nano-scale phenomena [3, 4]. A major difference of EIT from CIT is the adoption of an extended entropy function, which depends on not only the state variables  $\mathbf{z}(t)$ , but also their time/space derivatives (or flux variables) representing the transport processes. Except for this point, the formulation of EIT is very similar to that of CIT and will not be addressed further here.

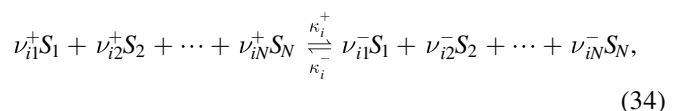
The intrinsic difference between EIT and CIT could be understood in a better way based on the large deviations principle. Actually, by studying the macroscopic limit of an  $\epsilon$ -dependent Langevin dynamics [29], it has been shown that the stationary large deviations rate functions of probability density  $p_\epsilon(\mathbf{x}, t)$  and joint probability density  $p_\epsilon(\mathbf{x}, \dot{\mathbf{x}}, t)$  turn out to be the flux-independent entropy function in CIT and flux-dependent entropy function in EIT, respectively. A summary of various thermodynamically compatible modeling approaches, including both conservative and dissipative dynamics, can be found in table 1 below.

## 4. Applications

In this section, we examine the GCS from numerical aspects. Through two specific examples—general chemical mass-action equations under detailed balance condition and the classical diffusion equation—the application procedures of various thermodynamic modeling approaches are demonstrated. In particular, the dissipation function in the GCS is illustrated numerically.

### 4.1. Chemical mass-action equations

For a spatially homogeneous chemical reaction system with  $N$  species and  $M$  reversible reactions



the governing equation for the concentration of the  $k$ 'th

**Table 1.** Summary of thermodynamically compatible modeling approaches.

Approach	Key elements	Governing equations
Conservative dynamics	Energy function $E(\mathbf{z})$ , antisymmetric matrix $\mathbf{A}(\mathbf{z})$	$\frac{d\mathbf{z}}{dt} = \mathbf{A}(\mathbf{z}) \cdot \nabla E(\mathbf{z})$
Hamiltonian dynamics	Hamiltonian function $H(\mathbf{z}, \boldsymbol{\xi})$	$\frac{d\mathbf{z}}{dt} = \frac{\partial H}{\partial \boldsymbol{\xi}}, \frac{d\boldsymbol{\xi}}{dt} = -\frac{\partial H}{\partial \mathbf{z}}$
Lagrangian dynamics	Lagrangian function $L(\mathbf{z}, \dot{\mathbf{z}})$	$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\mathbf{z}}} \right) - \frac{\partial L}{\partial \mathbf{z}} = \mathbf{0}$
Gradient flow	Free energy $F(\mathbf{z})$	$\frac{d\mathbf{z}}{dt} = -\nabla F(\mathbf{z})$
Dissipative dynamics	Free energy $F(\mathbf{z})$ , dissipation matrix $\mathbf{M}(\mathbf{z})$	$\frac{d\mathbf{z}}{dt} = -\mathbf{M}(\mathbf{z}) \cdot \nabla F(\mathbf{z})$
Generalized gradient systems	Free energy $F(\mathbf{z})$ , dissipative functions $\psi(\mathbf{z}, \dot{\mathbf{z}})$ , $\psi^*(\mathbf{z}, \boldsymbol{\xi})$	$\psi(\mathbf{z}, \dot{\mathbf{z}}) + \psi^*(\mathbf{z}, -\nabla F)$ $+ \langle \nabla F, \dot{\mathbf{z}} \rangle = 0$
Onsager variational principle	Free energy $F(\mathbf{z})$ , dissipative function $\psi(\mathbf{z}, \dot{\mathbf{z}})$	$\frac{\partial \mathcal{R}(\mathbf{z}, \dot{\mathbf{z}})}{\partial \dot{\mathbf{z}}} = \mathbf{0}$ , $\mathcal{R}(\mathbf{z}, \dot{\mathbf{z}}) = \psi(\mathbf{z}, \dot{\mathbf{z}}) + \langle \nabla F, \dot{\mathbf{z}} \rangle$
Energetic variational approach	Action $A_F(\mathbf{z}, \dot{\mathbf{z}})$ , dissipative function $\psi(\mathbf{z}, \dot{\mathbf{z}})$	$\frac{\partial A_F(\mathbf{z}, \dot{\mathbf{z}})}{\partial \dot{\mathbf{z}}} = \frac{\partial \psi(\mathbf{z}, \dot{\mathbf{z}})}{\partial \dot{\mathbf{z}}}$ , $A_F = \int_0^t (K - F) d\tau$

( $k = 1, \dots, N$ ) species reads

$$\frac{dc_k}{dt} = \sum_{i=1}^M (\nu_{ik}^- - \nu_{ik}^+) \left( \kappa_i^+ \prod_{j=1}^N c_j^{\nu_{ij}^+} - \kappa_i^- \prod_{j=1}^N c_j^{\nu_{ij}^-} \right), \quad (35)$$

based on the mass-action law. Here  $c_k(t) = [S_k](t)$  is the concentration for the  $k$ 'th species,  $\nu_{ik}^+$  and  $\nu_{ik}^-$  are stoichiometric coefficients for the  $k$ 'th species in the  $i$ 'th reaction, and  $\kappa_i^+$  and  $\kappa_i^-$  are the forward and backward reaction rate constants, respectively.

The above reaction system fulfills the detailed-balance condition if and only if there exists a positive steady state  $\mathbf{c}_* \equiv (c_1^*, \dots, c_N^*)^T > \mathbf{0}$  such that

$$\kappa_i^+ \prod_{j=1}^N c_j^{*\nu_{ij}^+} = \kappa_i^- \prod_{j=1}^N c_j^{*\nu_{ij}^-} \equiv K_i^*, \quad \forall i = 1, \dots, M. \quad (36)$$

Under the detailed-balance steady state  $\mathbf{c}_*$ , the forward reaction rate exactly equals the corresponding backward reaction rate for each reaction. Denoting the concentration vector  $\mathbf{c} \equiv (c_1, \dots, c_N)^T$ , we have the free energy or relative entropy

$$F(\mathbf{c}) = \sum_{i=1}^N (c_i \ln c_i - c_i \ln c_i^* - c_i + c_i^*). \quad (37)$$

*GCS for chemical mass-action equations.* Under the condition of detailed balance, the chemical mass-action equation (35) obeys the GCS [30] as

$$\frac{d\mathbf{c}}{dt} = \frac{\partial \psi^*(\mathbf{c}, \boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \Bigg|_{\boldsymbol{\xi} = -\nabla F}, \quad (38)$$

where the relative entropy is chosen as in equation (37), and the antisymmetric operator  $\mathbf{A}(\mathbf{c}) = \mathbf{0}$ . The dissipation function  $\psi^*(\mathbf{c}, \boldsymbol{\xi})$  is defined by choosing an arbitrary smooth

dissipation functional  $\psi_i$

$$\psi^*(\mathbf{c}, \boldsymbol{\xi}) = \sum_{i=1}^M Q_i(\mathbf{c}) \psi_i((\nu_i^+ - \nu_i^-) \cdot \boldsymbol{\xi}),$$

$$Q_i(\mathbf{c}) = K_i^* \frac{\frac{c_i^{\nu_i^-}}{c_*^{\nu_i^-}} - \frac{c_i^{\nu_i^+}}{c_*^{\nu_i^+}}}{\psi_i' \left( \ln \frac{c_i^{\nu_i^-}}{c_*^{\nu_i^-}} - \ln \frac{c_i^{\nu_i^+}}{c_*^{\nu_i^+}} \right)}. \quad (39)$$

Here, the vectors  $\nu_i^+ = (\nu_{i1}^+, \dots, \nu_{iN}^+)$ ,  $\nu_i^- = (\nu_{i1}^-, \dots, \nu_{iN}^-)$ , the function  $\psi_i: \mathbb{R} \rightarrow [0, \infty)$ ,  $\psi_i(0) = \psi_i'(0) = 0$ , and  $\psi_i'' > 0$ .

*GENERIC for chemical mass-action equations.* Under the condition of detailed balance, Yong showed that the mass-action equations could be cast into the GENERIC form [31]:

$$\frac{d\mathbf{c}}{dt} = -\mathbf{M}(\mathbf{c}) \nabla F, \quad \mathbf{M}(\mathbf{c}) = \sum_{i=1}^M e^{\sigma_i(\mathbf{c})} \boldsymbol{\Sigma}_i, \quad (40)$$

where the  $N \times N$  matrix  $\boldsymbol{\Sigma}_i \equiv [(\nu_{ik}^+ - \nu_{ik}^-)(\nu_{ij}^+ - \nu_{ij}^-)]$ , is symmetric, and the positive value

$$e^{\sigma_i(\mathbf{c})} = \left( \kappa_i^- \prod_{j=1}^N c_j^{\nu_{ij}^-} - \kappa_i^+ \prod_{j=1}^N c_j^{\nu_{ij}^+} \right) / \left( \ln \left( \kappa_i^- \prod_{j=1}^N c_j^{\nu_{ij}^-} \right) - \ln \left( \kappa_i^+ \prod_{j=1}^N c_j^{\nu_{ij}^+} \right) \right),$$

is determined by the mean-value theorem. It is straightforward to verify that  $\mathbf{M}(\mathbf{c}) = \sum_{i=1}^M e^{\sigma_i(\mathbf{c})} \boldsymbol{\Sigma}_i$  is a symmetric and positive semi-definite matrix, and the null space of  $\mathbf{M}(\mathbf{c})$  is independent of  $\mathbf{c}$ .

Notice that in this case, the relative entropy is also chosen as in equation (37). Comparing the GCS and the GENERIC of the chemical mass-action equations, we observe an analogous structure appearing in both function  $Q_i(\mathbf{c})$  in the GCS and  $e^{\sigma_i(\mathbf{c})}$  in the GENERIC, which also is the key for linking the kinetics and thermodynamics in both formulations.

*EVA for chemical mass-action equations.* Wang *et al* [32] showed that the chemical mass-action equation can be derived from the EVA under the condition of detailed balance. Ignoring the kinetic energy  $\mathcal{K}$  in the overdamped region, we choose the total energy of the system as  $E_{tot} = F(\mathbf{c})$  in equation (37). We denote the stoichiometric matrix  $\nu = [(\nu_{ij})] \in \mathbb{R}^{M \times N}$  with elements  $\nu_{ij} = \nu_{ij}^- - \nu_{ij}^+$ . In the context of differential geometry, the reaction trajectory  $\mathbf{R}(t)$  is introduced to account for the number of forward reactions that have occurred by time  $t$ ,

$$\mathbf{c}(\mathbf{R}) = \mathbf{c}^0 + \mathbf{R} \cdot \nu, \quad (41)$$

in which  $\mathbf{c}^0$  denotes the initial concentrations. It is noticeable that the reaction trajectory naturally preserves the stoichiometric constraints, especially the conservation laws.

Reformulating the free energy in terms of  $\mathbf{R}(t)$ , we have

$$F(\mathbf{c}(\mathbf{R})) = \sum_{i=1}^N (c_i \ln c_i - c_i \ln c_i^* - c_i + c_i^*), \quad (42)$$

where  $\mathbf{c}(\mathbf{R})$  is given in equation (41). On the other hand, the rate of energy dissipation is chosen as

$$\Delta_E(\mathbf{R}; \mathbf{R}_t) = \sum_{i=1}^M \partial_t R_i \ln \left( \frac{\partial_t R_i}{\eta_i(\mathbf{c}(\mathbf{R}))} + 1 \right), \quad (43)$$

where  $\eta_i(\mathbf{c}(\mathbf{R}))$  is a nonlinear function of  $\mathbf{c}$ . The free energy and the rate of energy dissipation function constitute an energy-dissipation law:

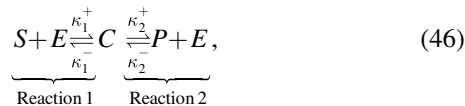
$$\frac{d}{dt} F(\mathbf{c}(\mathbf{R})) = -\Delta_E(\mathbf{R}; \mathbf{R}_t). \quad (44)$$

Considering the fact that  $dF/dt = \langle \delta F / \delta \mathbf{R}, \mathbf{R}_t \rangle$  and  $\delta F / \delta \mathbf{R} = \nu \ln(\mathbf{c}/\mathbf{c}^*)$ , the EVA actually implies

$$\nu \ln \frac{\mathbf{c}}{\mathbf{c}^*} = -\ln \left( \frac{\partial_t \mathbf{R}}{\eta(\mathbf{c}(\mathbf{R}))} + \mathbf{1} \right). \quad (45)$$

By choosing  $\eta_i(\mathbf{c}) = \kappa_i^- \prod_{j=1}^N c_j^{\nu_{ij}^-}$  and recalling the definition  $\kappa_i^+ / \kappa_i^- = \prod_{j=1}^N (c_j^*)^{\nu_{ij}^+}$ , it is straightforward to show that  $\partial_t R_i = \kappa_i^+ \prod_{j=1}^N c_j^{\nu_{ij}^+} - \kappa_i^- \prod_{j=1}^N c_j^{\nu_{ij}^-}$  for  $i = 1, \dots, M$ . This recovers the original mass-action equation in equation (35).

*Numerical results on Michaelis–Menten reactions.* We consider the Michaelis–Menten (MM) reactions based on the enzyme–substrate binding mechanism. There are four species, substrate  $S$ , enzyme  $E$ , complex  $C \equiv SE$ , and product  $P$ , which participate in two reversible reactions:



where  $\kappa_i^\pm > 0$  ( $i = 1, 2$ ) are reaction rate constants. We denote  $\mathbf{c}(t) \equiv (c_1, c_2, c_3, c_4)^\top = ([S], [E], [C], [P])^\top$ . The stoichiometric matrix becomes  $\nu = \begin{bmatrix} -1 & -1 & 1 & 0 \\ 0 & -1 & 1 & -1 \end{bmatrix}$ . Based on the two conservation laws of enzyme and substrate,  $[E] + [C] = [E]_{tot}$ ,  $[S] + [C] + [P] = [S]_{tot}$ , along with the detailed-balance condition, there exists a unique positive steady state,  $\mathbf{c}_* \equiv (c_1^*, c_2^*, c_3^*, c_4^*)^\top > 0$  for the MM reactions.

In figure 1, the dynamic and thermodynamic structures of the MM reactions are presented. As illustrated in figures 1(a), (b), the temporal concentrations tend to the equilibrium states with time, while the free energy function  $F(t)$  decreases monotonically, and its negative gradient  $f_d(t) \geq 0$  (inset) has a sudden decrease in a short period of time. For the MM reactions, the free energy  $F(\mathbf{c}) = \sum_{i=1}^4 (c_i \ln c_i - c_i \ln c_i^* - c_i + c_i^*)$ , which is rewritten as  $F([S], [E])$  by substitution of  $[C] = [E]_{tot} - [E]$  and  $[P] = [S]_{tot} - [E]_{tot} - [S] + [E]$  due to the conservation laws of enzyme and substrate. As shown in figure 1(c), the black solid line on the free energy surface, representing the trajectory from the given initial state, a point with a higher free energy to the global minimum. The dissipation function  $\psi^*([S], \xi_1)$  versus the concentration  $[S]$  and  $\xi_1$  is illustrated in figure 1(d). Here, we choose the quadric function in equation (39) as  $\psi_i(x) = x^2/2$  for all  $i = 1, 2$ . The concentration  $[E] \equiv [E]_{tot}/2$  is fixed, and the concentrations  $[C]$  and  $[P]$  are determined based on the conservation laws, while the remaining variables are set to zero,  $\xi_2 = \xi_3 = \xi_4 \equiv 0$  for simplicity.

#### 4.2. Diffusion equations

In the second example, we consider the classical diffusion system on a bounded smooth domain  $\Omega \subset \mathbb{R}^d$ . The governing equation for the density  $\rho = \rho(t, \mathbf{x})$  reads

$$\partial_t \rho(t, \mathbf{x}) = D \Delta_{\mathbf{x}} \rho, \quad (47)$$

where the constant  $D > 0$  is the diffusion coefficient. The relative entropy for this system is

$$F(\rho) = \int_{\Omega} (\rho \ln \rho - \rho) d\mathbf{x}. \quad (48)$$

*GCS for diffusion equations.* The classical diffusion equation (47) obeys the GCS [33] as

$$\frac{\partial \rho}{\partial t} = \frac{\partial \psi^*(\rho, \xi)}{\partial \xi} \Bigg|_{\xi = -\nabla_{\rho} F}, \quad (49)$$

where the dissipation function  $\psi^*(\rho, \xi)$  is defined as

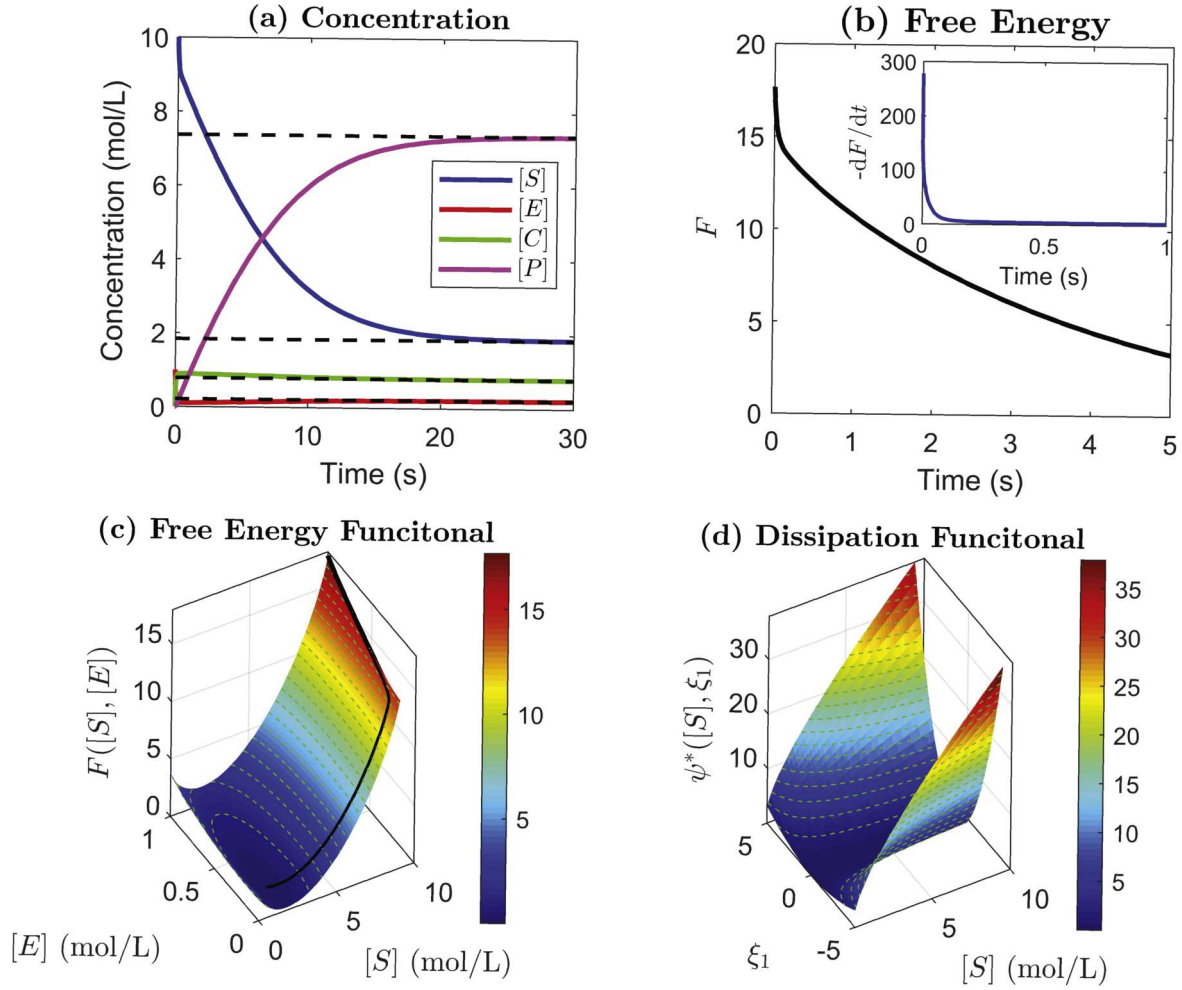
$$\psi^*(\rho, \xi) = \int_{\Omega} \frac{1}{2} D \rho |\nabla_{\mathbf{x}} \xi|^2 d\mathbf{x}. \quad (50)$$

In the above, the relative entropy is chosen as in equation (48), and the antisymmetric operator  $\mathbf{A}(\mathbf{c}) = \mathbf{0}$ . In fact, using the relations  $\frac{\partial \psi^*(\rho, \xi)}{\partial \xi} = -\nabla \cdot (D \rho \nabla \xi)$  and  $\nabla_{\rho} F = \ln \rho$ , we find  $\frac{\partial \psi^*(\rho, \xi)}{\partial \xi} \Big|_{\xi = -\nabla F} = \nabla \cdot (D \rho \nabla \ln \rho) = D \Delta_{\mathbf{x}} \rho$ .

*EVA for diffusion equations.* With respect to the continuity equation,

$$\partial_t \rho(t, \mathbf{x}) = -\nabla_{\mathbf{x}} \cdot (\rho \mathbf{v}), \quad (51)$$

where  $\mathbf{v} = \mathbf{v}(t, \mathbf{x})$  denotes the velocity. Wang *et al* [32] showed that the diffusion equation can be obtained from the EVA. We choose the total energy of the system as  $E_{tot} = F(\rho)$  in equation (48), and the rate of energy



**Figure 1.** Dynamic and thermodynamic structures of the MM reactions. (a) The concentrations versus time of the species  $S$ ,  $E$ ,  $C$ , and  $P$ ; the solid lines represent temporal states, while the dashed lines represent equilibrium states. (b) The free energy function  $F(t)$  and its negative gradient  $f_d(t)$  (inset) versus time. (c) The free energy surface  $F([S], [E])$  versus the concentrations  $[S]$  and  $[E]$ , while the black solid line represents the trajectory from the given initial state. (d) The dissipation function  $\psi^*([S], \xi_1)$  versus the concentration  $[S]$  and  $\xi_1$ . For all plots, the initial concentrations are  $([S], [E], [C], [P])|_{t=0} = (10, 1, 0, 0)$ , and the rate constants are  $(\kappa_1^+, \kappa_1^-, \kappa_2^+, \kappa_2^-) = (2, 1, 1, 0.5)$ .

dissipation as

$$\Delta_E(\rho; \mathbf{v}) = \int_{\Omega} D^{-1} \rho |\mathbf{v}|^2 dx. \quad (52)$$

Considering the fact that  $dF/dt = \langle \delta F / \delta \rho, \partial_t \rho \rangle$  and  $\delta F / \delta \rho = \ln \rho$ , the energy-dissipation law  $\frac{d}{dt} F(\rho) = -\Delta_E(\rho; \mathbf{v})$  actually implies

$$\begin{aligned} \langle \ln \rho, \partial_t \rho \rangle &= \langle \ln \rho, -\nabla_x \cdot (\rho \mathbf{v}) \rangle \\ &= \langle \rho \mathbf{v}, \nabla_x (\ln \rho) \rangle = \langle \rho \mathbf{v}, -D^{-1} \mathbf{v} \rangle, \end{aligned} \quad (53)$$

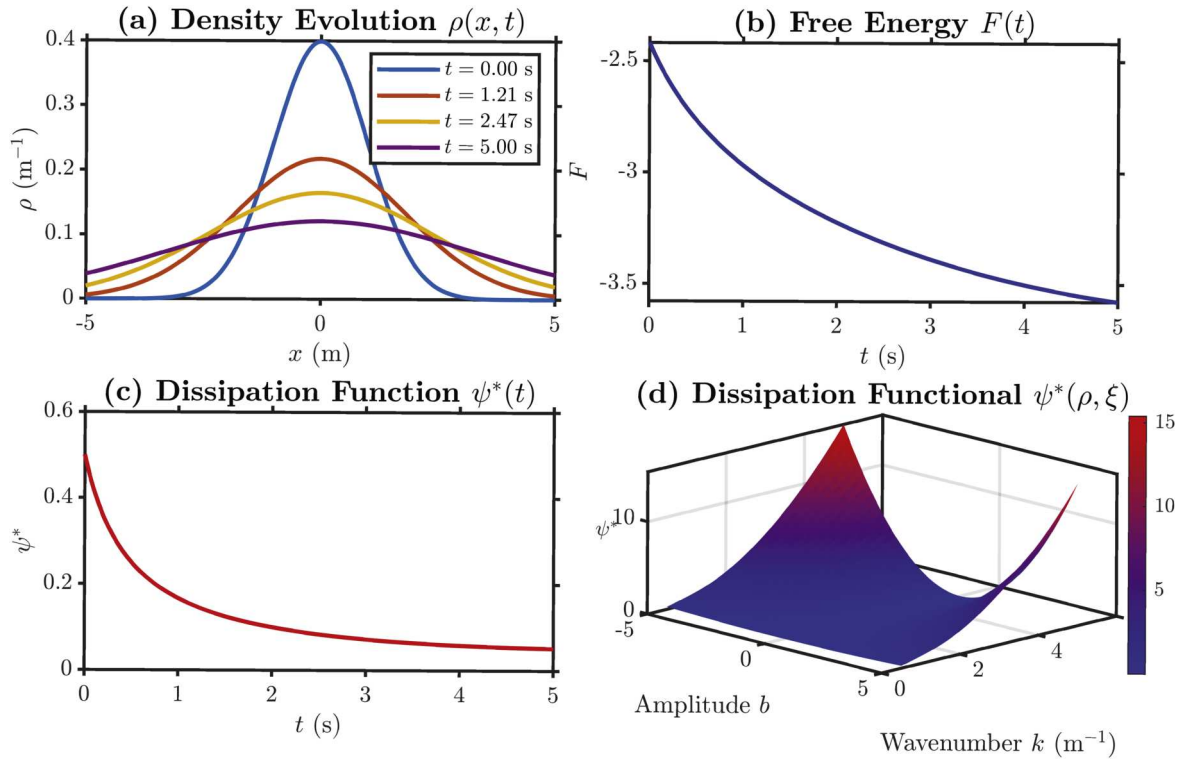
where in the second equality we used integration by parts and ignored the boundary terms. Based on the last equality, we can derive the constitutive relation of velocity  $\mathbf{v} = -D \nabla_x (\ln \rho)$ , which is recognized as Fick's law of diffusion. Direct substitution of Fick's law into the continuity equation recovers the original diffusion equation in equation (47).

*CIT for diffusion equations.* The classical diffusion equation can be constructed on the basis of classical irreversible thermodynamics, too. The starting point is also the continuity equation in equation (51). To find a proper

constitutive relation for the velocity  $\mathbf{v}$ , we turn to the entropy function  $S(\rho) = [U(\rho) - F(\rho)]/T = \int_{\Omega} (\rho u_0 - \rho \ln \rho + \rho) dx$ , where  $u_0$  denotes the specific internal energy. Recalling the Gibbs relation, we have

$$\begin{aligned} \frac{dS}{dt} &= \frac{1}{T} \frac{d}{dt} \int_{\Omega} (\rho u_0 - \rho \ln \rho + \rho) dx \\ &= \frac{1}{T} \int_{\Omega} \frac{\partial \rho}{\partial t} (u_0 - \ln \rho) dx \\ &= -\frac{1}{T} \int_{\Omega} \nabla_x \cdot (\rho \mathbf{v}) (u_0 - \ln \rho) dx \\ &= \frac{1}{T} \int_{\Omega} \rho \mathbf{v} \cdot \nabla_x (u_0 - \ln \rho) dx \\ &= \frac{1}{T} \int_{\Omega} \mathbf{v} \cdot \nabla_x \rho dx = \text{epr} \geq 0. \end{aligned}$$

In order to keep the non-negativeness of the entropy production rate, we can take Fick's law, which in turn gives the classical diffusion equation.



**Figure 2.** Dynamic and thermodynamic structures of the diffusion equations. (a) Density profiles  $\rho(t, x)$  at four different times. (b) Temporal evolution of the free energy function  $F(t) = \int (\rho \ln \rho - \rho) dx$ . (c) Temporal evolution of the dissipation function  $\psi^*(t) = \int \frac{1}{2} D \rho |\nabla \xi|^2 dx$  along the path  $\xi = -\nabla_\rho F = -\ln \rho$ . (d) Dissipation functional surface  $\psi^*(\rho, \xi)$  at fixed  $t = 2.47$  with  $\xi(x) = b \cdot \sin(2\pi kx/L)$ , illustrating its dependence on the amplitude  $b$  and wavenumber  $k$ . For all plots, the diffusion coefficient  $D = 1.0$ , spatial domain  $[-L, L]$  with  $L = 20$ .

*Numerical results on diffusion equations.* In figure 2, the dynamic and thermodynamic structures of the diffusion equations are presented. Figure 2(a) shows the temporal evolution of the density profile  $\rho(t, x)$ , which evolves from an initial Gaussian distribution throughout the spatial domain  $[-L, L]$  with  $L = 20$  at  $t = 0$  towards a uniform equilibrium state as  $t$  increases. Figures 2(b), (c) jointly demonstrate the thermodynamic consistency of the diffusion process. The free energy in (b) decreases monotonically. Meanwhile, the dissipation function  $\psi^*(t) = \int \frac{1}{2} D \rho |\nabla \xi|^2 dx$  in (c) along the path  $\xi = -\nabla_\rho F = -\ln \rho$  decays rapidly from its maximum at  $t = 0$  to zero as the system reaches equilibrium. Figure 2(d) illustrates the dissipation functional surface  $\psi^*(\rho, \xi)$  at fixed time  $t = 2.47$ , evaluated for trial functions  $\xi(x) = b \cdot \sin(2\pi kx/L)$ . The parabolic dependence on the amplitude  $b$  reflects the quadratic form of the dissipation functional, while the increasing curvature with the wavenumber  $k$  demonstrates the enhanced dissipation at smaller length scales.

## 5. Conclusion and discussion

In this paper, we make a thorough revision of the GCS, which includes the Hamiltonian dynamics and the gradient system as two specific cases. We highlight the elegant mathematical properties of the generalized gradient system, including its close relation to the large deviations principle,

reformulation in Wasserstein space, extension in metric spaces, as well as physical interpretation in the laws of thermodynamics. Most importantly, based on the GCS, we demonstrate a unification of various thermodynamic modeling approaches, including the GENERIC, OVP, the EVA and CIT, and suggest that they can lead to consistent results by proper formulation. Through our results, the intrinsic connections among different thermodynamic modeling approaches are clearly revealed, which sheds light on the construction of a unified theoretical framework for modeling various non-equilibrium thermodynamic processes in the future.

Though the GCS provides a very powerful framework with solid modern mathematical foundations, the classical physical systems (or models) that could be rewritten into GCS are still quite limited. A major difficulty is how to find a proper dissipation function and justify its mathematical properties as required. In addition, the separation of conservative and dissipative parts with respect to a new model is usually not straightforward. Another interesting point worthy of further exploration is the merits of GCS in data-driven modeling.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (12205135), Guangdong Basic and Applied Basic Research Foundation (2023A1515010157),

and the Natural Science Foundation of Fujian Province of China (2024J01212). LH acknowledges helpful discussions from Prof. Hong Qian and Bin Miao. The authors gratefully acknowledge the anonymous reviewers for their constructive suggestions, which have helped greatly in improving the clarity of this manuscript.

## Appendix A: Proof of the mathematical properties of the dissipation function

Now we are going to prove the mathematical properties of the dissipation function, which is defined through the large deviations principle stated in equation (13).

1.  $\psi^*(z, \xi)$  is convex with respect to the second argument  $\xi$ ;
2.  $\psi^*(z, \xi) \geq 0 \forall \xi \in T_z^*Z$ , and  $\psi^*(z, \mathbf{0}) = 0$ ;
3.  $\psi^*(z, \xi) = \psi(z, -\xi) \forall \xi \in T_z^*Z$ ,

**Proof.** 1. Since  $\psi^*(z, \xi) = H(z, \nabla U + \xi) - H(z, \nabla U)$ , and  $H(z, \xi)$  is the Legendre–Fenchel transformation of the Lagrangian function, the convexity of  $\psi^*(z, \xi)$  with respect to the second argument  $\xi$  is obvious.

2. Due to the time reversibility of the Lagrangian function, we see that  $\nabla U = \frac{\partial L(z, \dot{z})}{\partial \dot{z}}|_{\dot{z}=\mathbf{0}}$ , which means  $\nabla U$  is the minimizer of  $H(z, \xi)$  by duality. Therefore, we have  $\psi^*(z, \xi) \geq 0 \forall \xi \in T_z^*Z$ . In addition, the correctness of  $\psi^*(z, \mathbf{0}) = 0$  is obvious.

3. Actually, we can prove that the following statements are equivalent:

- (i)  $\psi^*(z, \xi) = \psi^*(z, -\xi) \forall z \in Z, \forall \xi \in T_z^*Z$ ;
- (ii)  $\psi(z, \zeta) = \psi(z, -\zeta) \forall z \in Z, \forall \zeta \in T_z Z$ ;
- (iii)  $L(z, \zeta) - L(z, -\zeta) = 2\langle \nabla U, \zeta \rangle \forall z \in Z, \forall \zeta \in T_z Z$ ;
- (iv)  $H(z, \nabla U + \xi) = H(z, \nabla U - \xi) \forall z \in Z, \forall \xi \in T_z^*Z$ .

The equivalence of the time symmetry between  $\psi^*$  and  $\psi$  is a direct consequence of their duality relation. From the expression of the Lagrangian in equation (15), it is easy to see that the time reversibility of the Lagrangian function implies the symmetry of time of  $\psi$ , and vice versa. From equation (13), the condition (i) is equivalent to the time reversibility of the Hamiltonian.

## Appendix B: Gradient flows in metric spaces

Let  $(\Omega, d)$  be a metric space, let  $F: \Omega \rightarrow [0, \infty]$  be a function, and let  $x: [0, T] \rightarrow \Omega$  be a locally absolutely continuous curve. Then the gradient flow of  $F(x(t))$  is defined through the following energy dissipation inequality (EDI) [34]

$$\int_0^t \left( \frac{1}{2} |\dot{x}(s)|^2 + \frac{1}{2} |\nabla_x F(x(s))|^2 \right) ds + F(x(t)) - F(x(0)) \leq 0,$$

for almost every  $0 < t < T$ . The metric derivative

$$|\dot{x}(t)| := \limsup_{s \rightarrow t} \frac{d(x(s), x(t))}{|s - t|}, \quad a.e. \ 0 < t < T,$$

and the norm of the gradient

$$|\nabla_x F(x(s))| := \limsup_{x \rightarrow \bar{x}} \frac{\max(0, F(\bar{x}) - F(x))}{d(x, \bar{x})},$$

are defined separately.

The relationship between EDI and gradient flows is straightforward. Considering

$$\begin{aligned} 0 &\leq \frac{1}{2} |\dot{x}(s) + \nabla_x F(x(s))|^2 \\ &= \frac{1}{2} |\dot{x}(s)|^2 + \frac{1}{2} |\nabla_x F(x(s))|^2 + \dot{x}(s) \cdot \nabla_x F(x(s)) \\ &= \frac{1}{2} |\dot{x}(s)|^2 + \frac{1}{2} |\nabla_x F(x(s))|^2 + \frac{dF(x(s))}{ds}, \end{aligned}$$

and integrating the formula above from 0 to  $t$ , we obtain

$$\int_0^t \left( \frac{1}{2} |\dot{x}(s)|^2 + \frac{1}{2} |\nabla_x F(x(s))|^2 \right) ds + F(x(t)) - F(x(0)) \geq 0.$$

Comparing the above equation with the EDI, we note that EDI holds for almost every  $t$  if and only if the curve  $x = x(t)$  solves the following gradient flow

$$\begin{cases} \dot{x}(t) = -\nabla_x F(x(t)), \\ x(0) = x_0. \end{cases}$$

## References

- [1] Demirel Y and Gerbaud V 2007 *Nonequilibrium Thermodynamics: Transport and Rate Processes in Physical, Chemical and Biological Systems* (Elsevier)
- [2] De Groot S R and Mazur P 1962 *Non-equilibrium Thermodynamics* (North-Holland Publishing Company)
- [3] Müller I and Ruggeri T 2013 *Rational Extended Thermodynamics* (Springer Science and Business Media)
- [4] Jou D, Casas-Vázquez J and Lebon G 2010 *Extended Irreversible Thermodynamics* (Springer)
- [5] Grmela M and Öttinger H C 1997 Dynamics and thermodynamics of complex fluids: I. Development of a general formalism *Phys. Rev. E* **56** 6620–32
- [6] Mielke A, Peletier M A and Renger D M 2014 On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion *Potential Analysis* **41** 1293–327
- [7] Doi M 2011 Onsager’s variational principle in soft matter *J. Phys.: Condens. Matter* **23** 284118
- [8] Liu C and Sun H 2009 On energetic variational approaches in modeling the nematic liquid crystal flows *Discrete Contin. Dyn. Sys.* **23** 455–75
- [9] Pelech P, Tuma K, Pavelka M, Sipka M and Sykora M 2020 On compatibility of the natural configuration framework with generic: derivation of anisotropic rate-type models arXiv:2010.06903
- [10] Mielke A, Patterson R I, Peletier M A and Michiel Renger D 2017 Non-equilibrium thermodynamical principles for chemical reactions with mass-action kinetics *SIAM J. Appl. Math.* **77** 1562–85
- [11] Freidlin M I and Wentzell A D 1998 *Random Perturbations* (Springer) 15–43
- [12] Adams S, Dirr N, Peletier M and Zimmer J 2013 Large deviations and gradient flows *Phil. Trans. R. Soc. A* **371** 20120341

- [13] Otto F 2001 The geometry of dissipative evolution equations: the porous medium equation *Commun. PDE* **26** 101–74
- [14] Santambrogio F 2015 *Optimal Transport for Applied Mathematicians* (Birkäuser)
- [15] Grmela M 1989 Hamiltonian mechanics of complex fluids *J. Phys. A: Math. Gen.* **22** 4375
- [16] Grmela M and Öttinger H C 1997 Dynamics and thermodynamics of complex fluids: II. Illustrations of a general formalism *Phys. Rev. E* **56** 6633–55
- [17] Öttinger H C 2018 Generic: review of successful applications and a challenge for the future arXiv:1810.08470
- [18] Onsager L 1931 Reciprocal relations in irreversible processes. I. *Phys. Rev.* **37** 405
- [19] Onsager L 1931 Reciprocal relations in irreversible processes. II. *Phys. Rev.* **38** 2265
- [20] Doi M 2013 *Soft Matter Physics* (Oxford University Press)
- [21] Doi M 2021 Onsager principle in polymer dynamics *Prog. Polym. Sci.* **112** 101339
- [22] Strutt J W 1871 Some general theorems relating to vibrations *Proc. London Math. Soc.* **1** 357–68
- [23] Onsager L and Machlup S 1953 Fluctuations and irreversible processes *Phys. Rev.* **91** 1505
- [24] Wang H, Qian T and Xu X 2021 Onsager’s variational principle in active soft matter *Soft Matter* **17** 3634–53
- [25] Liu C and Wu H 2019 An energetic variational approach for the Cahn–Hilliard equation with dynamic boundary condition: model derivation and mathematical analysis *Arch. Ration. Mech. Anal.* **233** 167–247
- [26] Wang Y, Zhang T-F and Liu C 2021 A two species micro–macro model of wormlike micellar solutions and its maximum entropy closure approximations: an energetic variational approach *J. Non-Newton. Fluid Mech.* **293** 104559
- [27] Wang Y and Liu C 2022 Some recent advances in energetic variational approaches *Entropy* **24** 721
- [28] Hong L and Qian H 2021 Stochastic dynamics, large deviation principle, and nonequilibrium thermodynamics *Phys. Rev. E* **104** 044113
- [29] Hong L and Qian H 2020 The statistical foundation of entropy in extended irreversible thermodynamics *J. Phys. A: Math. Theor.* **53** 425202
- [30] Maas J and Mielke A 2020 Modeling of chemical reaction systems with detailed balance using gradient structures *J. Stat. Phys.* **181** 2257–303
- [31] Yong W A 2012 Conservation-dissipation structure of chemical reaction systems *Phys. Rev. E* **86** 067101
- [32] Wang Y, Liu C, Liu P and Eisenberg B 2020 Field theory of reaction-diffusion: law of mass action with an energetic variational approach *Phys. Rev. E* **102** 062147
- [33] Liero M, Mielke A, Peletier M A and Renger D 2015 On microscopic origins of generalized gradient structures arXiv:1507.06322
- [34] Figalli A and Glaudo F 2023 *An Invitation to Optimal Transport, Wasserstein Distances, and Gradient Flows* (EMS Press)