Research Article

A. Mielke, M. A. Peletier, and D.R.M. Renger

A generalization of Onsager's reciprocity relations to gradient flows with nonlinear mobility

Abstract: Onsager's 1931 'reciprocity relations' result connects microscopic time-reversibility with a symmetry property of corresponding macroscopic evolution equations. Among the many consequences is a variational characterization of the macroscopic evolution equation as a gradient-flow, steepest-ascent, or maximal-entropy-production equation. Onsager's original theorem is limited to close-to-equilibrium situations, with a Gaussian invariant measure and a linear macroscopic evolution. In this paper we generalize this result beyond these limitations, and show how the microscopic time-reversibility leads to natural generalized symmetry conditions, which take the form of generalized gradient flows.

DOI: 10.1515/jnet-YYYY-XXXX

Keywords: gradient flows, generalized gradient flows, large deviations, symmetry, microscopic reversibility

1 Introduction

1.1 Onsager's reciprocity relations

In his two seminal papers in 1931, Lars Onsager showed how time-reversibility of a system implies certain symmetry properties of macroscopic observables of the system (Ons31; OM53). In modern mathematical terms the main result can be expressed as follows.

Theorem 1.1. *Let X_t be a Markov process in $\mathbb{R}^n$ with transition kernel $P_t(dx|x_0)$ and invariant measure $\mu(dx)$. Define the expectation $z_t(x_0)$ of X_t given that $X_0 = x_0$,*

$$z_t(x_0) = \mathbb{E}_{x_0} X_t = \int x P_t(dx|x_0).$$

Assume that

1. *μ is* reversible, *i.e., for all x, x_0, and $t > 0$, $\mu(dx_0)P_t(dx|x_0) = \mu(dx)P_t(dx_0|x)$;*
2. *μ is Gaussian with mean zero and covariance matrix G.*
3. *$t \mapsto z_t(x_0)$ satisfies the equation*

$$\dot{z}_t = -Az_t, \tag{1.1}$$

for some nonnegative $n \times n$ matrix A.

Then $M := AG$ is symmetric positive definite, and if we define $\mathcal{S}(x)$ by $\mu(dx) = \exp(\mathcal{S}(x))\,dx$, then equation (1.1) *can be written as*

$$\dot{z}_t = MD\mathcal{S}(z_t). \tag{1.2}$$

Here we write $D\mathcal{S}$ for the derivative of $\mathcal{S}$.

A. Mielke, D.R.M. Renger: *Weierstraß-Institut, Mohrenstraße 39, 10117 Berlin, Germany*
A. Mielke: *Institut für Mathematik, Humboldt-Universität zu Berlin, Rudower Chaussee 25, 12489 Berlin (Adlershof), Germany*
M. A. Peletier: *Department of Mathematics and Computer Science, and Institute for Complex Molecular Systems (ICMS), Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, the Netherlands*
M. A. Peletier: *Corresponding author; m.a.peletier@tue.nl*

Note that we purposefully avoid the use of thermodynamic terms such as temperature, pressure, or energy, in order to highlight that the result is essentially *mathematical* in nature (which becomes even more evident upon considering the proof as described in e.g. (dGM62, Sec. VII.4)). In the context of thermodynamics, the 'Markov process' typically is some stochastic process given by the macroscopic observables in question.

The symmetry property $M^T = M$ states equality of corresponding off-diagonal elements of M, which by equation (1.2) translates into equality of 'cross-effect coefficients'. These are known as the 'reciprocity relations', and are a corner stone of linear irreversible thermodynamics. In addition to the practical convenience of reducing the number of parameters, they also present a rare simplifying insight into the wild world of irreversible processes.

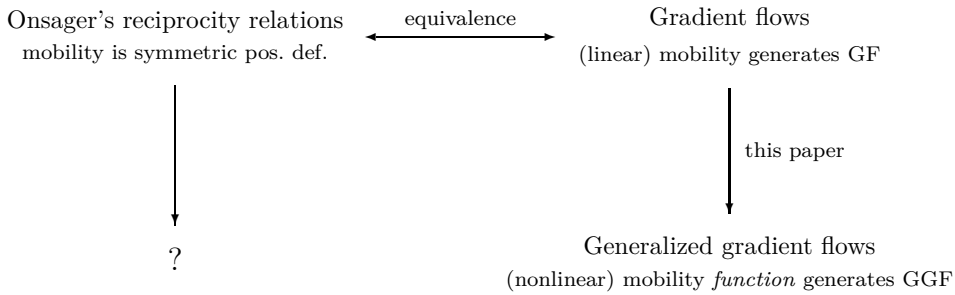
In its original form, however, Theorem 1.1 only applies to near-Gaussian fluctuations around a given stationary state, which translates into 'close to equilibrium'. For this reason, many generalizations have been proposed, e.g. (Cas45; Zie58; Gya70; HG82; GCdRC84; GC95; Gal96; MN07; Sei12; RZ15).

Despite these advances, there is still a major challenge. In Onsager's argument the macroscopic symmetry property arises from the microscopic reversibility, and it is exactly this connection that gives Onsager symmetry its force. This connection is generalized in our paper (MPR14); in the current paper we present this mathematical micro-to-macro connection from a more thermodynamic perspective.

The central observation that allows us to move forward is that the property 'symmetry of a matrix' is intrinsically linked to the linear nature of a matrix—the fact that a matrix is a linear operator. This linear nature is an obstacle to generalization to nonlinear behaviour. We therefore generalize Onsager's relations in three steps:

I. We first remark that symmetry and positive definiteness of the mobility M are equivalent to the property that equation (1.2) is a *gradient flow*;
II. We next generalize the concept of gradient flows (with linear mobilities) to *generalized gradient flows*, in which the mobility may be a nonlinear map;
III. We finally show that under more general conditions than Onsager's, the macroscopic evolution is such a generalized gradient flow.

This final point is the generalization of Onsager's relations: instead of guaranteeing symmetry of some linear operator, which is equivalent to the macroscopic equation being a gradient flow, we guarantee under more general conditions that the macroscopic equation is a *generalized* gradient flow. We argue in this paper that this generalized gradient-flow property is the natural generalization of Onsager symmetry, from the point of view of microscopic reversibility.



We walk through these three steps one by one in the sections that follow.

Remark 1.2. One should pay attention to the fact that in probability theory reversibility means detailed balance, whereas in thermodynamics, reversibility means that quantities such as entropy or free energy remain constant along evolutions. In fact, probabilistic *reversibility* often leads to thermodynamic *irreversibility*, as is the case here. □

Remark 1.3. The results that we discuss in this paper are described in mathematical terms and in full detail in (MPR14). In this paper we focus on the consequences of these results for thermodynamics. □

2 Step I: Definiteness is equivalent to gradient structure

We use the language of geometry, in which the *states* z are elements of a manifold $\mathcal{Z}$, and at each $z \in \mathcal{Z}$ there is a *tangent plane* $T_z\mathcal{Z}$ consisting of all vectors tangent (thermodynamic fluxes) to $\mathcal{Z}$ at z. If $t \mapsto z_t$ is a smooth curve along the manifold $\mathcal{Z}$, then the time derivative $\dot{z}_t$ is a tangent vector, i.e. $\dot{z}_t \in T_{z_t}\mathcal{Z}$. Dual to the tangent plane $T_z\mathcal{Z}$ at any $z \in \mathcal{Z}$ is the *cotangent plane* $T_z^*\mathcal{Z}$, whose elements (driving forces) are linear functionals on tangents, i.e. each $\xi \in T_z^*\mathcal{Z}$ is a linear mapping from $T_z\mathcal{Z}$ to $\mathbb{R}$. We write $\langle s, \xi \rangle$ for the dual pairing between $\xi \in T_z^*\mathcal{Z}$ and $s \in T_z\mathcal{Z}$, and we write $T\mathcal{Z}$ and $T^*\mathcal{Z}$ for the collection of all tangent and cotangent planes.

Definition 2.1. *We call the triple* $(\mathcal{Z}, \mathcal{S}, M)$ *a* gradient system *if* $\mathcal{S} : \mathcal{Z} \to \mathbb{R}$ *and for each* $z \in \mathcal{Z}$, $M(z)$ *is a symmetric positive definite linear mapping from* $T_z^*\mathcal{Z}$ *to* $T_z\mathcal{Z}$.

A curve $z : [0, T] \to \mathcal{Z}$ *is then called a* solution of the gradient flow *if it solves the differential equation*

$$\dot{z}_t = M\big(z_t\big)D\mathcal{S}\big(z_t\big) \qquad \textit{or equivalently} \qquad D\mathcal{S}\big(z_t\big) = M(z_t)^{-1}\dot{z}_t. \tag{2.1}$$

Although the equation (2.1) is also meaningful for more general linear mappings M, we purposefully require symmetry and positive definiteness in order to call (2.1) a gradient flow, for a number of reasons. First, the positive-definiteness forces $\mathcal{S}$, which in many cases one can interpret as an entropy or a negative free energy, to *increase* along solutions, since

$$\frac{d}{dt}\mathcal{S}\big(z_t\big) = \langle D\mathcal{S}\big(z_t\big), \dot{z}_t\rangle = \langle D\mathcal{S}\big(z_t\big), M\big(z_t\big)D\mathcal{S}\big(z_t\big)\rangle \geq 0. \tag{2.2}$$

Secondly, positive-definiteness implies invertibility of $M(z)$ (there are generalizations possible where M is only semidefinite). The third reason lies in the variational characterization of $\dot{z}$ that we discuss below. If $M(z)$ and $M(z)^{-1}$ are symmetric and positive definite, then they define norms $\frac{1}{2}\xi M(z)\xi$ and $\frac{1}{2}s^T M(z)^{-1}s$ on cotangents $\xi \in T_z^*\mathcal{Z}$ and tangents $s \in T_z\mathcal{Z}$ respectively.

Using the norm on the tangent space, the evolution can also be given a variational characterization, already recognized by Onsager, as

a curve $t \mapsto z_t$ is a solution of the gradient flow if and only for each t and z_t,

$$\dot{z}_t \text{ solves the maximization problem } \max_{s \in T_z\mathcal{Z}} D\mathcal{S}\big(z_t\big)^T s - \frac{1}{2}s^T M(z)^{-1}s.$$

Since the first term in the minimised expression above can be interpreted as the instantaneous production of $\mathcal{S}$, this evolution is often interpreted as 'steepest ascent' of $\mathcal{S}$ (e.g. (Ber87; MS06; Ber14)). This variational characterization depends critically on the symmetry of M, and this is why the name 'gradient system' is reserved for symmetric operators M.

Note that in such a *gradient flow* the mobility operator $M = M(z)$ may depend in any way on the position z, but the dependence on the argument $\xi = D\mathcal{S}(z)$ is linear: $\xi \mapsto M(z)\xi$ is a linear map. This is different from the case of nonlinear mobilities M, which are often written as $\xi \mapsto M(z, \xi)\xi$, in which this map depends nonlinearly on ξ. The *generalized* gradient-flow concept that we now introduce allows for nonlinear dependence on ξ.

3 Step II: Generalized gradient systems

We generalize to systems with nonlinear mobility M by noting that a gradient flow can alternatively be formulated by defining, for $z \in \mathcal{Z}, s \in T_z\mathcal{Z}$ and $\xi \in T_z^*\mathcal{Z}$,

$$\Psi(z, s) := \frac{1}{2}\langle M(z)^{-1}s, s\rangle \qquad \text{and its convex dual} \qquad \Psi^*(z, \xi) := \frac{1}{2}\langle \xi, M(z)\xi\rangle. \tag{3.1}$$

Here the convex duality is in the sense of Legendre-Fenchel transforms,

$$\Psi^*(z,\xi) = \sup_{s\in T_z\mathcal{Z}} \langle \xi, s\rangle - \Psi(z,s), \qquad \text{and} \qquad \Psi(z,s) = \sup_{\xi\in T_z^*\mathcal{Z}} \langle \xi, s\rangle - \Psi^*(z,\xi). \tag{3.2}$$

Equation (2.1) can then also be written as

$$D_s\Psi(z_t,\dot z_t) = D\mathcal{S}(z_t) \qquad \text{or equivalently} \qquad \dot z_t = D_\xi\Psi^*\big(z_t, D\mathcal{S}(z_t)\big), \tag{3.3}$$

which, by Legendre-Fenchel theory, is equivalent to requiring

$$\Psi(z_t,\dot z_t) + \Psi^*\big(z_t, D\mathcal{S}(z_t)\big) - \langle D\mathcal{S}(z_t), \dot z_t\rangle = 0. \tag{3.4}$$

Generalized gradient systems are again of the form (3.3) and (3.4), but where Ψ and Ψ^* need not be of the quadratic form (3.1). Such formulations have a long history of study in both the thermodynamic (Ede72; Grm93; GÖ97; Grm10; Grm12) and the mathematical literature (DGMT80; DG93; CV90; LS95; MRS08; Mie11a; Mie11b), under various different names, such as Dissipation potentials, Entropy-dissipation structures, Ψ-Ψ^*-structures, and many others. Here, Ψ and Ψ^* need to satisfy a number of conditions in order to yield a meaningful generalized formulation; these are summarized in the following definition.

Definition 3.1. *We call the triple* $(\mathcal{Z},\mathcal{S},\Psi)$ *a* generalized gradient system *if* $\mathcal{S}:\mathcal{Z}\to\mathbb{R}$, $\Psi: T\mathcal{Z}\to\mathbb{R}$, *and for all* $z\in\mathcal{Z}$*:*

1. $\Psi(z,\cdot)$ *is convex in the second argument,*
2. $\min\Psi(z,\cdot) = 0$*, and*
3. $\Psi(z,0) = 0$.

In addition, Ψ *is called* symmetric *if* $\Psi(z,s) = \Psi(z,-s)$ *for all* $(z,s)\in T\mathcal{Z}$.

A curve $z:[0,T]\to\mathcal{Z}$ *satisfying* (3.3) *or* (3.4) *is then called a solution of the* generalized gradient flow *induced by the generalized gradient system* $(\mathcal{Z},\mathcal{S},\Psi)$.

With the duality (3.2) we have the following equivalences for Conditions 2 and 3,

$$\min\Psi(z,\cdot) = 0 \qquad \Longleftrightarrow \qquad \Psi^*(z,0) = 0, \tag{3.5a}$$

$$\Psi(z,0) = 0 \qquad \Longleftrightarrow \qquad \min\Psi^*(z,\cdot) = 0, \tag{3.5b}$$

so that both $\Psi\geq 0$ and $\Psi^*\geq 0$. This is a central part of the definition; it implies that for curves satisfying (3.4), we have the nonlinear analogue of (2.2):

$$\frac{d}{dt}\mathcal{S}(z_t) = \langle D\mathcal{S}(z_t),\dot z_t\rangle = \Psi(z_t,\dot z_t) + \Psi^*(z_t, D\mathcal{S}(z_t)) \geq 0.$$

Remark 3.2. In a generalized gradient flow the role of the mobility is played by the (potentially nonlinear) map $\xi\mapsto D_\xi\Psi^*(z,\xi)$. Although this allows for nonlinear mobilities, the requirement that $D_\xi\Psi^*$ is a derivative means that not all nonlinear mobilities $M(z,\xi)\xi$ can be written as $M(z,\xi)\xi = D_\xi\Psi^*(z,\xi)$ for some Ψ^*; see (HS13) for a discussion. □

4 Step III: Reversibility and large deviations lead to generalized gradient structures

We now come to our main point. We generalize Onsager's result by connecting reversibility with *generalized* gradient flows, but in a more general context than that of Theorem 1.1. Instead of a fixed system we choose a system with a parameter n, and consider the limit $n\to\infty$. This allows us to treat a very general class of systems using the theory of large deviations.

Let Z_t^n be a sequence of stochastic processes with transition kernel P_t^n and invariant measure μ^n. Typically, Z_t^n models a system with microscopic fluctuations, where n is a large quantity like the number of particles in the system. Assume that

(A) P_t^n and μ^n satisfy the same reversibility condition as in Theorem 1.1, i.e. $\mu^n(dx_0)P_t^n(dx|x_0) = \mu^n(dx)P_t^n(dx_0|x)$;

(B) The equilibrium distributions μ^n satisfy a large-deviation principle with rate function $\mathcal{S}$, i.e. if Y^n are a random variables with distribution μ^n, then

$$\mathrm{Prob}(Y^n \approx y) \underset{n\to\infty}{\sim} \exp\big[n\mathcal{S}(y)\big];$$

(C) For all $T > 0$, Z^n satisfies a large-deviation principle of the form

$$\mathrm{Prob}_{\mu^n}\Big(Z_t^n\big|_{t\in[0,T]} \approx z\big|_{t\in[0,T]}\Big) \underset{n\to\infty}{\sim} \exp\big[n\left(\mathcal{S}(z_0) - I_T(z)\right)\big], \quad I_T(z) := \int_0^T \mathscr{L}(z_t, \dot z_t)\,dt, \quad (4.1)$$

where the subscript μ^n indicates that the process starts at a position drawn from μ^n.

These three assumptions are natural generalizations of the conditions of Onsager's Theorem 1.1. Condition (B) above is a natural generalization of the requirement that μ is Gaussian (Theorem 1.1, condition 2). Note that since probabilities are bounded by one, Condition (B) implies that $\mathcal{S} \le 0$, which can be seen as the large-deviation counterpart of the normalization of probability measures; the most probable behaviour in equilibrium is given by the maximizer(s) of $\mathcal{S}$ with value zero.

Condition (C) above is the natural extension of the assumption on the linear evolution of macroscopic quantities in Theorem 1.1, condition 3, but this requires a little more explanation. The typical case is that $\mathscr{L}$ is what we call an *L-function*:

Definition 4.1. *We call $\mathscr{L}$ an L-function*[1] *if*

1. *$\mathscr{L} \ge 0$,*
2. *$\mathscr{L}(z, \cdot)$ is convex for all z,*
3. *$\mathscr{L}$ induces an evolution equation $\dot z_t = \mathcal{A}_{\mathscr{L}}(z_t)$, in the sense that for all (z, s),*

$$\mathscr{L}(z, s) = 0 \qquad \Longleftrightarrow \qquad s = \mathcal{A}_{\mathscr{L}}(z). \quad (4.2)$$

Again, the fact that $\min_s \mathscr{L}(z, s) = 0$ follows from the boundedness of probabilities. The convexity is a standard consequence of the lower-semicontinuity given by large-deviation theory, but it will be important below. The fact that $\mathscr{L}(z, \cdot)$ has a single minimum, however, is a strong requirement, which is satisfied if in the limit $n \to \infty$ the behaviour of Z_t^n becomes deterministic (a form of the law of large numbers). The limiting macroscopic equation is then $\dot z_t = \mathcal{A}_{\mathscr{L}}(z_t)$.

A very large class of systems satisfies the assumptions above: diffusion-type equations arising from stochastic lattice models, for instance (KOV89), or from many-particle Brownian motion (KO90), including convection and mean-field interaction (DG87; FK06). Chemical reactions are another important example, and we describe this in more detail in the next section. The method of Feng and Kurtz (FK06, Sec. 8.6.1.2) provides an algorithm to derive such large-deviation principles in great generality.

The following theorem is our main result.

Theorem 4.2. *Assume conditions (A)–(C) above. If $\mathscr{L}$ is an L-function, then it generates a generalized gradient system $(\mathcal{Z}, \frac{1}{2}\mathcal{S}, \Psi)$ with symmetric potential Ψ, i.e. the macroscopic evolution $\dot z_t = \mathcal{A}_{\mathscr{L}}(z_t)$ (see (4.2)) can be written as the generalized gradient flow $\dot z_t = D\Psi^*\big(z_t, \frac{1}{2}D\mathcal{S}(z_t)\big)$.*

Proof. The proof consists of two steps:

1. The reversibility implies a symmetry relation for $\mathscr{L}$ and $\mathcal{S}$,

$$\mathscr{L}(z, s) - \mathscr{L}(z, -s) = -\langle D\mathcal{S}(z), s\rangle. \quad (4.3)$$

1 The letter 'L' in 'L-function' refers to the word *Lagrangian*, a name often used for expressions of the form of I_T in (4.1). We thank the anonymous reviewer for pointing out that the name 'L-function' has also been used to indicate *Lyapunov* functions (see e.g. (Ber86)), which is different concept (but note that Theorem 4.2 implies that an L-function $\mathscr{L}$ with certain properties *generates* a Lyapunov function $\mathcal{S}$).

2. This symmetry implies that there exists a pair of potentials Ψ, Ψ^*, satisfying the requirements of Definition 3.1, such that

$$\mathscr{L}(z,s) = \Psi(z,s) + \Psi^*\big(z, \tfrac{1}{2}D\mathcal{S}(z)\big) - \langle \tfrac{1}{2}D\mathcal{S}(z), s\rangle. \tag{4.4}$$

Part 1 is a well-known consequence of reversibility in the context of large deviations (Mae99; MRVM00; Sei12), and its proof runs as follows. For an arbitrary $(z,s) \in T\mathcal{Z}$, take any curve for which $(z_0, \dot{z}_0) = (z,s)$. From the reversibility follows the exact characterization, for all $T > 0$,

$$\mathrm{Prob}_{\mu^n}\Big(Z_t^n\big|_{t\in[0,T]} \approx z\big|_{t\in[0,T]}\Big) = \mathrm{Prob}_{\mu^n}\Big(Z_t^n\big|_{t\in[0,T]} \approx \sigma_T z\big|_{t\in[0,T]}\Big)$$

where σ_T is the time-reversal operator $(\sigma_T z)_t = z_{T-t}$. By (4.1) this implies

$$\begin{aligned}\mathcal{S}(z_0) - \int_0^T \mathscr{L}(z_t, \dot{z}_t)\,dt &= \mathcal{S}(z_T) - \int_0^T \mathscr{L}(z_{T-t}, -\dot{z}_{T-t})\,dt \\ &= \mathcal{S}(z_T) - \int_0^T \mathscr{L}(z_s, -\dot{z}_s)\,ds.\end{aligned}$$

Upon dividing by T and taking the limit $T \to 0$ we find (4.3).

The proof of part 2 follows from a manipulation of Legendre transforms; we refer to (MPR14, Th. 2.1) for the details. □

5 Example: chemical reactions

A large number of (generalized) gradient-flow formulations of macroscopic evolution equations can be connected to large-deviation principles of their underlying stochastic processes, in the way that we describe above (ADPZ11; DLR13; ADPZ13; DPZ13; BP14; MPR14; LMPR). Here we discuss a single example, that of chemical reactions.

We consider chemical reactions of I species $X_1, \dots, X_I$ of the form

$$\alpha_{1j}X_1 + \cdots + \alpha_{Ij}X_I \underset{r_j^-}{\overset{r_j^+}{\rightleftharpoons}} \beta_{1j}X_1 + \cdots + \beta_{Ij}X_I, \qquad j = 1, \dots, J,$$

under constant temperature and pressure.

Two natural models of this system are the Chemical Master Equation (e.g. (Gil00)) and the Reaction-Rate Equation (e.g. (ÉT89)). The Chemical Master Equation (CME) describes a continuous-time, discrete-space stochastic process in terms of numbers of particles of each type; the Reaction-Rate Equation (RRE) is a deterministic equation for concentrations of species:

$$\dot{\boldsymbol{c}}_t = \sum_{j=1}^{J} \big(r_j^+(\boldsymbol{c}_t) - r_j^-(\boldsymbol{c}_t)\big)(\boldsymbol{\beta}_j - \boldsymbol{\alpha}_j). \tag{5.1}$$

Under reasonable conditions on the kinetics, in the limit of large volume and under well-mixed conditions, the CME becomes deterministic, and its limit is given by the RRE (Kur73).

If the system satisfies both detailed balance and mass action, then (5.1) admits a stationary set of concentrations $\overline{\boldsymbol{c}} = (\overline{c}_1, \dots, \overline{c}_I)$ and the net rate of reaction j can be written as

$$r_j(\boldsymbol{c}) := r_j^+(\boldsymbol{c}) - r_j^-(\boldsymbol{c}) = k_j\big[\boldsymbol{u}^{\boldsymbol{\alpha}j} - \boldsymbol{u}^{\boldsymbol{\beta}j}\big] \qquad \text{where} \qquad u_i := \frac{c_i}{\overline{c}_i},$$

and we write

$$\boldsymbol{u}^{\boldsymbol{\alpha}j} := u_1^{\alpha_{1j}} \dots u_I^{\alpha_{Ij}} \qquad \text{and} \quad \boldsymbol{u}^{\boldsymbol{\beta}j} := u_1^{\beta 1j} \dots u_I^{\beta Ij}.$$

Note that typically the evolution takes place in a submanifold, generated by the particular values of the quantities that the evolution preserves. Therefore $\overline{c}$ is not the only stationary state—each stoichiometic subspace, defined by a particular value of all conserved quantities, contains exactly one stationary state.

The generalized gradient structure $(\mathcal{Z}, \frac{1}{2}\mathcal{S}, \Psi)$ for (5.1) is given by the driving functional

$$\mathcal{S}(\boldsymbol{c}) := -\sum_{i=1}^{I} \eta\Big(\frac{c_i}{\overline{c}_i}\Big)\overline{c}_i = -\sum_{i=1}^{I} \eta(u_i)\overline{c}_i, \qquad \eta(s) := s\ln s - s + 1, \tag{5.2}$$

and the exponential dissipation potential

$$\Psi_j^*(\boldsymbol{c}, \boldsymbol{\xi}) = 2k_j \boldsymbol{u}^{\frac{1}{2}\boldsymbol{\alpha}_j + \frac{1}{2}\boldsymbol{\beta}_j}\big(\cosh\big((\boldsymbol{\alpha}_j - \boldsymbol{\beta}_j)\cdot\boldsymbol{\xi}\big) - 1\big), \qquad \Psi^*(\boldsymbol{c}, \boldsymbol{\xi}) = \sum_{j=1}^{J} \Psi_j^*(\boldsymbol{c}, \boldsymbol{\xi}). \tag{5.3}$$

The functional $\mathcal{S}$ can be interpreted as a dimensionless negative free energy, for instance by noting that the derivative of $\mathcal{S}(\boldsymbol{c})$ with respect to a concentration c_i equals

$$\partial_{c_i}\mathcal{S}(\boldsymbol{c}) = -\eta'\Big(\frac{c_i}{\overline{c}_i}\Big) = -\log c_i + \log \overline{c}_i,$$

which is equal, up to constant factors and a minus sign, to the common expression for the chemical potential $\mu_i = \mu_i^{\mathrm{std}} + RT\log c_i$. Nonquadratic potentials such as (5.3) have been discussed by Grmela (Grm12) and similar expressions appear in the earlier literature (Fei72).

The restriction to the submanifold mentioned above is implicitly implemented by Ψ^*: since $D\Psi_j^*$ is parallel to $\boldsymbol{\alpha}_j - \boldsymbol{\beta}_j$, the range of $D\Psi^*$ is the span of the vectors $\{\boldsymbol{\alpha}_j - \boldsymbol{\beta}_j\}_{j=1}^J$.

For monomolecular reactions we prove in (MPR14) that this generalized gradient system is indeed generated by the large deviations of the corresponding Chemical Master Equation, which is the discrete stochastic process in terms of numbers of particles of each species. In the forthcoming paper (MPPR15) we show that the same is true for general mass-action systems with detailed balance. This shows that this generalized-gradient-flow structure of chemical reactions has its origins in the large-deviations behaviour of the underlying stochastic systems. The same is true for various other processes, such as diffusion (ADPZ13), heat conduction (PRV14), passive transport across membranes (GM13), and thermally activated motion (BP14), and we conjecture that this is actually a very general phenomenon.

6 Discussion

The strength of Onsager's theorem lies in the combination of its broad scope (all processes in a certain class) with a rigorous statement (the macroscopic equations have a certain structure). We generalize this theorem by enlargening the class to all sequences of processes with a large-deviation principle; our assertion, also rigorous, is an appropriate generalization of Onsager's symmetry statement, which coincides with the original symmetry in the case of linear processes.

Incidentally, this work shows how large-deviation theory, originally developed to understand mathematically such physical concepts as free energy and entropy, provides us with the tools also to understand dynamical generalizations of these concepts. In this way it appears to be the natural link between the stochastic microscopic description and the deterministic macroscopic one.

Interestingly, the large-deviation viewpoint of this paper explains the appearance of nonquadratic dissipation potentials in thermodynamics: when the underlying stochastic process has non-Gaussian fluctuations, then the large-deviation rate functional has non-quadratic dependence on the generalized velocity, and the dissipation potential in the generalized gradient flow is non-quadratic. The non-quadratic nature therefore can be traced back to non-Gaussian fluctuations.

It would be very natural to consider systems with varying time-reversal parity, following Casimir (Cas45), but at this moment this is only understood for a single example (DPZ13).

Acknowledgement. The authors are grateful to Gian Paolo Beretta, Miroslav Grmela, Markus Hütter, and Hans-Christian Öttinger for their comments on an early version of this paper. AM and DRMR were partially supported by DFG via SFB 1114 (projects C5 and C8); MAP was supported by the NWO VICI grant 639.033.008.

References

[ADPZ11] S. Adams, N. Dirr, M. A. Peletier, and J. Zimmer. From a large-deviations principle to the Wasserstein gradient flow: A new micro-macro passage. *Communications in Mathematical Physics*, 307:791–815, 2011.

[ADPZ13] S. Adams, N. Dirr, M. A. Peletier, and J. Zimmer. Large deviations and gradient flows. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(2005):20120341, 2013.

[Ber86] G. P. Beretta. A theorem on Lyapunov stability for dynamical systems and a conjecture on a property of entropy. *Journal of Mathematical Physics*, 27(1):305–308, 1986.

[Ber87] G. P. Beretta. Steepest entropy ascent in quantum thermodynamics. In *The Physics of Phase Space Nonlinear Dynamics and Chaos Geometric Quantization, and Wigner Function*, pages 441–443. Springer, 1987.

[Ber14] G. P. Beretta. Steepest entropy ascent model for far-nonequilibrium thermodynamics: Unified implementation of the maximum entropy production principle. *Physical Review E*, 90(4):042113, 2014.

[BP14] G. A. Bonaschi and M. A. Peletier. Quadratic and rate-independent limits for a large-deviations functional. *arXiv preprint arXiv:1409.4350*, 2014.

[Cas45] H. B. G. Casimir. On onsager's principle of microscopic reversibility. *Reviews of Modern Physics*, 17(2-3):343, 1945.

[CV90] P. Colli and A. Visintin. On a class of doubly nonlinear evolution equations. *Communications in Partial Differential Equations*, 15(5):737–756, 1990.

[DG87] D. A. Dawson and J. Gartner. Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics*, 20(4):247–308, 1987.

[DG93] E. De Giorgi. New problems on minimizing movements. In *Boundary value problems for partial differential equations and applications*, volume 29 of *RMA Res. Notes Appl. Math.*, pages 81–98. Masson, Paris, 1993.

[dGM62] S. R. de Groot and P. Mazur. *Non-Equilibrium Thermodynamics*. North-Holland, Amsterdam, 1962.

[DGMT80] E. De Giorgi, A. Marino, and M. Tosques. Problems of evolution in metric spaces and maximal decreasing curve. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)*, 68(3):180–187, 1980.

[DLR13] M. H. Duong, V. Laschos, and D. R. M. Renger. Wasserstein gradient flows from large deviations of many-particle limits. *ESAIM: Control, Optimisation and Calculus of Variations*, E-first, 2013.

[DPZ13] M. H. Duong, M. A. Peletier, and J. Zimmer. GENERIC formalism of a Vlasov-Fokker-Planck equation and connection to large-deviation principles. *Nonlinearity*, 26(2951-2971), 2013.

[Ede72] D. G. B. Edelen. A nonlinear Onsager theory of irreversibility. *International Journal of Engineering Science*, 10(6):481–490, 1972.

[ÉT89] P. Érdi and J. Tóth. *Mathematical Models of Chemical Reactions: Theory and Applications of Deterministic and Stochastic Models*. Manchester University Press, 1989.

[Fei72] M. Feinberg. On chemical kinetics of a certain class. *Archive for Rational Mechanics and Analysis*, 46(1):1–41, 1972.

[FK06] J. Feng and T. G. Kurtz. *Large deviations for stochastic processes*, volume 131 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2006.

[Gal96] G. Gallavotti. Extension of Onsager's reciprocity to large fields and the chaotic hypothesis. *Physical Review Letters*, 77(21):4334, 1996.

[GC95] G. Gallavotti and E. Cohen. Dynamical ensembles in nonequilibrium statistical mechanics. *Physical Review Letters*, 74(14):2694, 1995.

[GCdRC84] L. García-Colín and J. del Rio-Correa. Further generalization of the onsager reciprocity theorem. *Physical Review A*, 30(6):3314, 1984.

[Gil00] D. T. Gillespie. The chemical Langevin equation. *The Journal of Chemical Physics*, 113:297, 2000.

[GM13] A. Glitzky and A. Mielke. A gradient structure for systems coupling reaction–diffusion effects in bulk and interfaces. *Zeitschrift für angewandte Mathematik und Physik*, 64(1):29–52, 2013.

[GÖ97] M. Grmela and H. C. Öttinger. Dynamics and thermodynamics of complex fluids. i. development of a general formalism. *Physical Review E*, 56(6):6620, 1997.

[Grm93] M. Grmela. Weakly nonlocal hydrodynamics. *Physical Review E*, 47(1):351, 1993.

[Grm10] M. Grmela. Multiscale equilibrium and nonequilibrium thermodynamics in chemical engineering. *Advances in Chemical Engineering*, 39:75–129, 2010.

[Grm12] M. Grmela. Fluctuations in extended mass-action-law dynamics. *Physica D: Nonlinear Phenomena*, 241(10):976–986, 2012.

[Gya70] I. Gyarmati. *Non-equilibrium thermodynamics*. Springer, 1970.

[HG82] J. Hurley and C. Garrod. Generalization of the onsager reciprocity theorem. *Physical Review Letters*, 48(23):1575, 1982.

[HS13] M. Hütter and B. Svendsen. Quasi-linear versus potential-based formulations of force–flux relations and the generic for irreversible processes: comparisons and examples. *Continuum Mechanics and Thermodynamics*, 25(6):803–816, 2013.

[KO90] C. Kipnis and S. Olla. Large deviations from the hydrodynamical limit for a system of independent Brownian particles. *Stochastics and stochastics reports*, 33(1-2):17–25, 1990.

[KOV89] C. Kipnis, S. Olla, and S. R. S. Varadhan. Hydrodynamics and large deviation for simple exclusion processes. *Communications on Pure and Applied Mathematics*, 42(2):115–137, 1989.

[Kur73] T. G. Kurtz. Convergence of sequences of semigroups of nonlinear operators with an application to gas kinetics. *Trans. Amer. Math. Soc.*, 186:259–272, 1973.

[LMPR] M. Liero, A. Mielke, M. A. Peletier, and D. R. M. Renger. On microscopic origins of generalized gradient structures. In preparation.

[LS95] S. Luckhaus and T. Sturzenhecker. Implicit time discretization for the mean curvature flow equation. *Calculus of Variations and Partial Differential Equations*, 3(2):253–271, 1995.

[Mae99] C. Maes. The fluctuation theorem as a Gibbs property. *Journal of Statistical Physics*, 95(1/2):367–392, 1999.

[Mie11a] A. Mielke. Formulation of thermoelastic dissipative material behavior using GENERIC. *Continuum Mechanics and Thermodynamics*, 23(3):233–256, 2011.

[Mie11b] A. Mielke. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 24:1329–1346, 2011.

[MN07] C. Maes and K. Netočnỳ. Minimum entropy production principle from a dynamical fluctuation law. *Journal of mathematical physics*, 48(5):053306, 2007.

[MPPR15] A. Mielke, R. Patterson, M. Peletier, and D. Renger. Non-equilibrium thermodynamical principles for chemical reactions with mass-action kinetics. WIAS preprint No. 2165, 2015.

[MPR14] A. Mielke, M. A. Peletier, and D. R. M. Renger. On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion. *Potential Analysis*, 41(4):1293–1327, 2014.

[MRS08] A. Mielke, R. Rossi, and G. Savaré. A metric approach to a class of doubly nonlinear evolution equations and applications. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, VII(5):97–169, 2008.

[MRVM00] C. Maes, F. Redig, and A. Van Moffaert. On the definition of entropy production, via examples. *Journal of Mathematical Physics*, 41(3):1528–1554, 2000.

[MS06] L. Martyushev and V. Seleznev. Maximum entropy production principle in physics, chemistry and biology. *Physics reports*, 426(1):1–45, 2006.

[OM53] L. Onsager and S. Machlup. Fluctuations anf irreversible processes. *Physical Review*, 91(6):1505–1512, 1953.

[Ons31] L. Onsager. Reciprocal relations in irreversible processes I & II. *Physical Review*, 37:405–426 and 38:2265–2279, 1931.

[PRV14] M. A. Peletier, F. Redig, and K. Vafayi. Large deviations in stochastic heat-conduction processes provide a gradient-flow structure for heat conduction. *Journal of Mathematical Physics*, 55(9):093301, 2014.

[RZ15] C. Reina and J. Zimmer. Entropy production and the geometry of dissipative evolution equations. *arXiv preprint arXiv:1507.01014*, 2015.

[Sei12] U. Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics*, 75(12):126001, 2012.

[Zie58] H. Ziegler. An attempt to generalize onsager's principle, and its significance for rheological problems. *Zeitschrift für angewandte Mathematik und Physik ZAMP*, 9(5-6):748–763, 1958.