Fluctuation symmetry leads to GENERIC equations with non-quadratic dissipation

Richard C. Kraaij\textsuperscript{1}, Alexandre Lazarescu\textsuperscript{2}, Christian Maes\textsuperscript{3}, Mark A. Peletier\textsuperscript{4}

September 10, 2018

Abstract

We develop a formalism to discuss the properties of GENERIC systems in terms of corresponding Hamiltonians that appear in the characterization of large-deviation limits. We demonstrate how the GENERIC structure naturally arises from a certain symmetry in the Hamiltonian, which extends earlier work that has connected the large-deviation behaviour of reversible stochastic processes to the gradient-flow structure of their deterministic limit. Natural examples of application include particle systems with inertia.

Contents

1 Introduction 2

2 Variational characterizations of evolution equations 4
   2.1 Gradient flows 5
   2.2 GENERIC systems 6
   2.3 Pre GENERIC 8

3 Microscopic origins of variational evolution and fluctuation symmetries 9
   3.1 Microscopic origins of macroscopic equations and large deviations of processes 9
   3.2 Gradient flows and fluctuation symmetry 10
   3.3 Pre GENERIC and fluctuation symmetry around a vector field 12
   3.4 Returning to stochastic processes: Pre GENERIC from large deviations 14

4 From pre GENERIC to GENERIC 15
   4.1 Adding an energy variable for the macroscopic dynamics 16
   4.2 Transforming GENERIC in the setting with a natural energy 18

5 Two related examples: the Andersen thermostat and the Vlasov Fokker Planck equation 21
   5.1 The macroscopic dynamics and pre GENERIC 21
   5.2 Microscopic model and its description in terms of macroscopic variables 25
   5.3 Stationary and path space large deviations 27
1 Introduction

For a long time there has been a ‘folk theorem’ in the probability community that states the following: if a stochastic particle system is reversible, and if the particle system converges to a deterministic limit as the number of particles tends to infinity, then this deterministic limit is a gradient flow. Such a property has wide implications: it implies the presence of a Lyapunov function, it identifies a geometry of the underlying space that is generated by the noise, it constrains the type of evolution that the limit equation can show, and it gives essential insight in both the limit equation and the connection between the limit equation and the underlying stochastic process.

For linear stochastic differential equations with Gaussian noise this ‘folk theorem’ was shown by Onsager and Machlup [15], and this property has also been identified for various specific examples [2, 9, 10, 20]. All these works, however, are limited to the case of ‘quadratic’ noise (we explain this below, cf. Remark 2.3), and exclude for instance jump processes. The theorem was recently proven under assumptions of smoothness, but otherwise in full generality, covering both quadratic and non-quadratic noise, by Mielke, Peletier, and Renger [13]. Their result states that for any sequence of reversible stochastic processes with a deterministic limit and corresponding large-deviation principle, the large-deviation rate functional gives rise to a gradient-flow structure for the deterministic limit.

Since that result is central to this paper we briefly describe the global argument. Pathwise large-deviation principles of sequences of stochastic processes, reversible or not, can be expressed in so-called Hamiltonians and their duals, called Lagrangians (see Section 3.1). If the processes are reversible, then the Hamiltonian and the Lagrangian possess a certain symmetry. This symmetry implies that the Lagrangian can be written in a form that can be recognised as a generalization of the classical gradient-flow concept.

The aim of this paper is to do the same for a class of non-reversible processes: we connect a symmetry property of a class of non-reversible stochastic processes to a variational-evolutionary structure for a corresponding class of deterministic limiting equations. This class, known as GENERIC, consists of evolution equations for an unknown $z \in Z$ of the type

$$\dot{z} = L(z)dE(z) + \partial_\xi \Psi^*(z, -\frac{1}{2}dS(z)).$$

Here $E, S : Z \to \mathbb{R}$ are two functionals with derivatives $dS$ and $dE$, for each $z \in Z$ the operator $L(z)$ maps the derivative $dE(z)$ to a velocity vector, and $(z, \xi) \mapsto \Psi^*(z, \xi)$ is a non-negative functional. We discuss this structure in more detail in Section 2.2.

Gradient flows are special cases of GENERIC equations, in which the term $L(z)dE(z)$ is absent, and our point of view in this paper is that the appearance of the additional term $L(z)dE(z)$ in GENERIC can be connected to the breaking of the reversibility symmetry of the underlying stochastic processes. While we can not make this statement in full generality, we do make a number of steps in that direction, which form the novelty of this paper:
A class of energy preserving generalised-reversible stochastic processes

\[ H = H_1 + H_2 \]

\[ H_1(z,ξ) = (L(z)dE(z),ξ) \]

\[ L(z)dS(z) = 0 \]

Corollary 3.7

zero-cost flow is GENERIC

Corollary 4.4

1-1 embedding, Corollary 4.3

1-1 embedding, Theorem 3.6

1-1 embedding, Theorem 4.1

Proposition 3.8

H is reversible with respect to S

Prop. & Thm. 2.1

zero-cost flow is pre GENERIC

generalization

1-1 embedding, Proposition 3.7

generalization

generalization

Figure 1 – This diagram illustrates the global structure of this paper; see the main text for the discussion. Many of the terms and notation are only defined in later sections.

A. As in the case of gradient flows, a central role is played by the Hamiltonians that arise in the large-deviation behaviour. The first step is to develop a formalism to express GENERIC systems in terms of Hamiltonians.

B. Using this formalism, and its interpretation in terms of large-deviation principles, we connect symmetry properties of the underlying stochastic processes with a notion we call pre-GENERIC.

C. We show that pre-GENERIC can always be extended to GENERIC by a trivial addition of an auxiliary energy variable, showing that pre-GENERIC captures the relevant structure of GENERIC.

D. We illustrate these ideas on two examples involving particles with inertia. One is the Vlasov-Fokker-Planck system; the other is known as the Andersen thermostat and combines deterministic drift with a jump process in momentum space.

Figure 1 illustrates these steps in a graphical manner. The bottom row is the existing theory of [13], which applies to reversible stochastic processes and gradient flows; the top two rows constitute the generalization that we construct in this paper.

Item A above corresponds to the arrows connecting the middle and right columns. Theorem 3.6 and Corollary 3.7 reformulate GENERIC and the new concept ‘pre-GENERIC’ in terms of Hamiltonians \( H \) and their properties, thus generalizing Theorem 2.1 of [13] to certain generalised-reversible Hamiltonians.

Item B appears as the corresponding arrows on the left, Proposition 3.8 and Corollary 4.4, which identify classes of stochastic processes that generate Hamiltonians with the properties as given in the middle column.

Item C corresponds to the arrows that connect the middle to the top row, Theorem 4.1 and Corollary 4.4.

These results are developed in the paper in the following way. In Section 2 we introduce gradient flows, GENERIC systems, and pre-GENERIC systems. In Section 3 we recall the emergence of gradient-flow structures from reversible stochastic processes as described in [13], and generalize this to certain generalised-reversible systems. Section 4 is devoted to the step
from ‘pre-GENERIC’ to ‘full’ GENERIC and we show there how each pre-GENERIC system can be converted without loss of generality into a full GENERIC system. Items A, B and C are treated in Sections 3 and 4.

Finally, in Section 5 we discuss two examples, the Vlasov-Fokker-Planck equation and the Andersen thermostat [1] (item D). The corresponding microscopic systems are similar in that they both combine inertial effects with dissipative effects, leading to pre-GENERIC and GENERIC macroscopic systems. While the Vlasov-Fokker-Planck equation has quadratic dissipation and was treated in [3], the Andersen thermostat is a jump process and therefore has non-quadratic dissipation. We comment in detail on the consequences of this difference.

This paper is accompanied by the paper [11] which regards the topic from a physical point of view and in which we treat a greater class of examples.

Finally, we want to mention that the results and proofs in Sections 2, 3 and 4 are carried out for smooth finite-dimensional manifolds. We believe, however, that most results carry over to a more general context. Indeed the examples that are considered in Section 5 are based in an infinite dimensional setting.

In addition, in Definition 2.1, we restrict ourselves to strictly convex and continuously differentiable dissipation potentials. We believe that convexity suffices.

We have made these simplifying assumptions to achieve full rigor for the results in Sections 2, 3 and 4, while on the other hand giving a full treatment of the various structural connections mentioned above.

Acknowledgement The authors gratefully acknowledge Manh Hong Duong, who pointed us towards the fluctuation symmetry (3.5) in an early stage of the work on this paper.

R.K. was supported by the Deutsche Forschungsgemeinschaft (DFG) via RTG 2131 High-dimensional Phenomena in Probability—Fluctuations and Discontinuity. A.L. was supported by the Interuniversity Attraction Pole - Phase VII/18 (Dynamics, Geometry and Statistical Physics) at the KU Leuven and the AFR PDR 2014-2 Grant No. 9202381 at the University of Luxembourg. M.A.P. was partially supported by NWO VICI grant 639.033.008.

2 Variational characterizations of evolution equations

The evolution equations of this paper all are members of the class of ‘variational evolutions’, generalizations of the classical gradient-flow, or steepest-descent, evolution in $\mathbb{R}^n$. These generalizations all have in common that a central role is played by one or more driving functionals and one or more duality maps. Since the role of the duality map is more easily understood in a geometric language, we use a formal geometric language throughout.

We consider a smooth manifold $Z$, with tangent space $TZ = \bigcup_{z \in Z} T_zZ = \bigcup_{z \in Z} \{z\} \times T_zZ$ and cotangent space $T^*Z = \bigcup_{z \in Z} T^*_zZ$. We will say that $V : Z \to TZ$ is a vector-field on $Z$ if $V$ is continuous and $V(z) \in \{z\} \times T_zZ$. Without loss of generality, we will only consider the second component and write $V(z) \in T_zZ$.

For a function $F \in C_b(Z)$, we write $C^1(Z)$ the continuously differentiable functions, i.e. if for each element $z \in Z$ there is an appropriate derivative $dF(z) \in T^*_zZ$ and the map $z \mapsto dF(z)$ is continuous.

A continuous-time evolution equation is a differential equation that describes the evolution of a quantity $z$ in a state space (or manifold) $Z$, which at its most abstract level can be written
\[ \dot{z} = F(z), \]  

(2.1)

where \( F \) is a vector field on \( Z \). Every evolution equation is of the form \( (2.1) \); variational evolution equations are a subclass described by properties of \( F \). Two types are central to this paper: gradient flows and GENERIC systems.

2.1 Gradient flows

Let \( \Psi : T Z \to [0, \infty] \) and let \( \Psi^* : T^* Z \to [0, \infty] \) be Legendre duals:

\[
\Psi^*(z, \xi) = \sup_{v \in T_z Z} \{ \langle v, \xi \rangle - \Psi(z, v) \}, \quad \Psi(z, v) = \sup_{\xi \in T^*_z Z} \{ \langle v, \xi \rangle - \Psi^*(z, \xi) \}.
\]

**Definition 2.1.** We say that a Legendre pair \( (\Psi, \Psi^*) \) are dissipation potentials if \( \Psi, \Psi^* \) are strictly convex, continuously differentiable and

\[
\Psi(z, 0) = \Psi^*(z, 0) = 0.
\]

(2.2)

We say that \( \Psi \) and \( \Psi^* \) are symmetric if \( \Psi(z, v) = \Psi(z, -v) \) for all \((z, v) \in T Z\) and if \( \Psi^*(z, \xi) = \Psi^*(z, -\xi) \) for all \((z, \xi) \in T^* Z\).

Dissipation potentials satisfy \( 0 = \min_{v \in T_z Z} \Psi(z, v) = \min_{\xi \in T^*_z Z} \Psi^*(z, \xi) \) (see [13, Sec. 1.2]).

Let \( S \in C^1(S) \) be some functional. By Legendre duality, we find for all \((z, v) \in T Z\) that

\[
\Psi(z, v) + \Psi^*(z, -\frac{1}{2} dS(z)) + \frac{1}{2} \langle F(z), dS(z) \rangle \geq 0.
\]

A gradient flow with respect to dissipation potentials \( (\Psi, \Psi^*) \) is a flow \( \dot{z} = F(z) \) such that

\[
\Psi(z, F(z)) + \Psi^*(z, -\frac{1}{2} dS(z)) + \frac{1}{2} \langle F(z), dS(z) \rangle = 0
\]

for all \( z \in Z \).

Clearly, the functional \( S \) is a Lyapunov function for the gradient flow as

\[
\frac{d}{dt} S(z(t)) = \langle F(z), dS(z) \rangle = -2 \left( \Psi(z, F(z)) + \Psi^*(z, -\frac{1}{2} dS(z)) \right) \leq 0
\]

by the non-negativity of \( (\Psi, \Psi^*) \).

By the properties of Legendre duality, we find the equivalent characterizations

\[
\dot{z} = F(z) \text{ is an } (S, \Psi, \Psi^*)\text{-gradient flow } \iff \dot{z} = \partial_z \Psi^*(z, -\frac{1}{2} dS(z)) \quad (2.3a)
\]

\[
\iff -\frac{1}{2} dS(z) = \partial_{\xi} \Psi(z, \dot{z}). \quad (2.3b)
\]

Note that \( \partial_z \Psi(z, \cdot) \) and \( \partial_{\xi} \Psi^*(z, \cdot) \) are each other’s inverse (i.e., \( \partial_z \Psi(z, \partial_{\xi} \Psi^*(z, \xi)) = \xi \) and vice versa). Also note that they are duality maps: they map the tangent plane \( T_z Z \) at \( z \)
to the cotangent plane \( T^*_z \mathbb{Z} \) at \( z \), in the case of \( \partial_v \Psi(z, \cdot) \), and vice versa in the case of \( \partial_\xi \Psi^*(z, \cdot) \). These duality maps are a central ingredient of variational evolution, both from a mathematical and from a modelling point of view. Mathematically, they are necessary to convert the cotangent vector \( \frac{1}{2} dS(z) \) into a tangent vector \( F(z) \) that can be equated to \( \dot{z} \); since \( dS(z) \) and \( \dot{z} \) live in different spaces, the gradient-flow concept is not well defined without such a duality map. From a modelling point of view, on the other hand, the duality maps are force-to-velocity conversions: if \( S \) is interpreted as an energy, then the derivative \( dS(z) \) is a generalized force; the gradient-flow equation (2.3a) describes how this force leads to the movement described by \( \dot{z} \), through the force-to-velocity, cotangent-to-tangent map \( \partial_\xi \Psi^*(z, \cdot) \).

**Remark 2.3.** An important subclass of gradient flows are those with dissipation potentials \((\Psi, \Psi^*)\) that are quadratic functions of their second arguments, i.e. both \( v \mapsto \Psi(z, v) \) and \( \xi \mapsto \Psi^*(z, \xi) \) are quadratic functions of \( v \) and \( \xi \) respectively. We refer to these as quadratic gradient flows. In that case, the operators \( M(z) : T^*_z \mathbb{Z} \to T_z \mathbb{Z} \) defined by \( \xi \mapsto M(z)(\xi) := \partial_\xi \Psi^*(z, \xi) \) are linear. In addition, they satisfy the symmetry property \( \langle M(z)(\xi), \eta \rangle = \langle M(z)(\eta), \xi \rangle \) for all \( \eta, \xi \in T^*_z \mathbb{Z} \); if \( M \) is a matrix, this corresponds to the symmetry of \( M \). Onsager termed this symmetry property the ‘reciprocal relations’ in his seminal papers in 1931 [14].

**Example 2.4.** To facilitate the exposition later on, we will be using a ‘running example’ of a very simple, quadratic-dissipation system. To start this example, we let \( Z = \mathbb{R}^n \), \( S : \mathbb{Z} \to \mathbb{R} \) some smooth function, and \( \Psi^*(z, \xi) := \frac{1}{2} \| \xi \|^2_M = \frac{1}{2} \xi^T M \xi \), where \( M \in \mathbb{R}^{n \times n} \) is a fixed, symmetric, positive semi-definite matrix.

Given these choices, the gradient-flow equation generated by \( Z \), \( S \), and \( \Psi^* \) is the classical gradient flow equation

\[
\dot{z} = -\frac{1}{2} M dS(z).
\]  

(2.4)

### 2.2 GENERIC systems

In addition to the well-studied gradient-flow systems, a second class of well studied systems are systems of GENERIC type (Generalized Equation for Non-Equilibrium Reversible-Irreversible Coupling \([7, 16, 18]\)). Such systems combine in an orthogonal way a gradient flow and a Hamiltonian flow. ‘Classical’ GENERIC is formulated in terms of the symmetric operator \( M \) \([7]\); here we consider a generalized version of GENERIC, where the symmetric matrix \( M \) has been replaced by the gradient of the convex functional \( \Psi^* \) \([5, 7, 12]\).

Formally, our definition of a GENERIC equation implies an evolution equation of the form

\[
\dot{z} = L(z) dE(z) + \partial_\xi \Psi^* \left( z, -\frac{1}{2} dS(z) \right).
\]  

(2.5)

The definition that we now choose is intended to allow for non-differential dissipation potentials.

**Definition 2.5.** A GENERIC equation with respect to two functionals \( S, E \in C^1(\mathbb{Z}) \), dissipation potentials \((\Psi, \Psi^*)\) and duality maps \( L \) is given by

\[
\Psi(z, \dot{z} - L(z) dE(z)) + \Psi^* \left( z, -\frac{1}{2} dS(z) \right) + \frac{1}{2} \langle \dot{z}, dS(z) \rangle = 0
\]  

(2.6)

where
(a) \((\Psi, \Psi^*)\) are symmetric dissipation potentials;
(b) \(L = \{L(z)\}_{z \in Z}\) is a family of operators \(L(z) : T^*_z Z \to T_z Z\) that are ‘anti-symmetric’: for every \(z \in Z\) and \(\xi, \eta \in T^*_z Z\) we have \(\langle L(z)\xi, \eta \rangle = -\langle L(z)\eta, \xi \rangle\), and the bracket \(\{F, G\} := \langle dF, LdG\rangle\), (applied to ‘smooth’ functions \(F_i : Z \to \mathbb{R}\)) satisfies the Jacobi identity:
\[
\{\{F_1, F_2\}, F_3\} + \{\{F_3, F_1\}, F_2\} + \{\{F_2, F_3\}, F_1\} = 0.
\]
(c) \(E, S, L, \Psi^*\) satisfy the degeneracy conditions
\[
L(z) dS(z) = 0, \quad \Psi^* (z, \xi + \alpha dE(z)) = \Psi^* (z, \xi) \forall \alpha. \tag{2.7}
\]

The conditions in Definition \ref{de:b} are generalizations of the original conditions on GENERIC as formulated by Grmela and Öttinger \cite{8, 17, 19}. In brief, the operator \(L\) is interpreted as defining a Poisson bracket \(\{\cdot, \cdot\}\), such that the equation \(\dot{z} = L(z)dE(z)\) defines a Hamiltonian system with Hamiltonian \(E\). The joint equation (2.5) then combines this Hamiltonian dynamics with a gradient-flow dynamics determined by \(S\) and \(\Psi^*\).

The degeneracy conditions impose constraints on how these two dynamics combine: they characterize a property of ‘non-interaction’ \cite{12} between the two types of dynamics. This can be recognized for instance in the evolution of the functionals \(E\) and \(S\). These functionals are routinely interpreted as an energy and an entropy, and the properties of a GENERIC equation underwrite this: by the degeneracy conditions, we find that along the flow
\[
\frac{dE}{dt} = \langle dE, LdE\rangle + \langle dE, \partial_z \Psi^* \left(z, -\frac{1}{2}dS(z)\right)\rangle = \langle dE, LdE\rangle + \partial_{\xi} \Psi^* \left(z, -\frac{1}{2}dS(z) + \alpha dE(z)\right) = 0,
\]
whereas the non-negativity of \((\Psi, \Psi^*)\) and the GENERIC equation
\[
\frac{dS}{dt} = \langle \dot{z}, dS(z)\rangle = -2 \left(\Psi(z, \dot{z} - L(z)dE(z)) + \Psi^* \left(z, -\frac{1}{2}dS(z)\right)\right) \leq 0. \tag{2.8}
\]

Therefore \(E\) is conserved and \(S\) decreases along a solution of (2.6). (We should emphasize that the direction of change of \(S\) is the opposite of the usual GENERIC convention: \(S\) decreases along the flow.)

Thus, a GENERIC equation characterizes a relaxation to equilibrium in terms of a dynamics that consists of two parts; one is dissipative, and is the gradient flow of an ‘entropy’ \(S\), whereas the other is Hamiltonian, and conserves the ‘energy’ \(E\).

**Example 2.6.** To continue the example, the gradient-flow equation \cite{2.3} can be converted into a simple GENERIC system by adding a conservative term
\[
\dot{z} = LdE(z) - \frac{1}{2}MdS(z), \tag{2.9}
\]
where we introduced a constant antisymmetric matrix \(L \in \mathbb{R}^{n \times n}\) and an energy function \(E \in C^1(Z)\). The second degeneracy condition in (2.7) reduces to \(MdE = 0\), as we see from an explicit calculation: for all \(\xi \in \mathbb{R}^n\),
\[
0 = \partial_{\xi} \Psi^* \left(z, \xi + \alpha dE(z)\right) \bigg|_{\alpha = 0} = \partial_{\xi} \frac{1}{2} \|\xi + \alpha dE(z)\|_3^2 \bigg|_{\alpha = 0} = \xi^T MdE(z).
\]
Therefore we find the two degeneracy conditions \(LdS = 0\) and \(MdE = 0\).
Example 2.7. We can make our ‘running example’ even more explicit. A natural example of equation (2.9) is a Hamiltonian system with friction; here $L_d E$ represents the Hamiltonian dynamics, and $-\frac{1}{2} M dS$ the dissipative dynamics associated with friction.

This example however cannot be done with a constant matrix $M$, and therefore we will consider a position-dependent matrix $M(z)$. See [12, Sec. 3.1] for a more extensive discussion of this example.

Let $q \in \mathbb{R}^d$ and $p \in \mathbb{R}^d$ be canonical variables of the Hamiltonian system, with Hamiltonian $H(q, p) = \frac{p^2}{2m} + V(q)$. Upon adding friction to the system, $H$ is not conserved; the thermodynamic interpretation is that mechanical energy (represented by $H$) is converted into heat. Therefore we add an ‘internal energy’ variable $e \in \mathbb{R}$ that captures this heat; we discuss this addition in more detail in Section 4. The state space is then $\mathbb{R}^{d+d+1}$, with variables $z = (q, p, e)$, and the total energy is $E(q, p, e) = H(q, p) + e = \frac{p^2}{2m} + V(q) + e$.

The Poisson operator $L$ simply describes the symplectic structure of the Hamiltonian system,

$$L = \begin{pmatrix} 0 & I & 0 \\ -I & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{(d+d+1)\times(d+d+1)}.$$

The dissipative structure is given by

$$S(q, p, e) := -e, \quad M(q, p, e) := 2\gamma \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -p/m \\ 0 & -p/m & p^2/m^2 \end{pmatrix},$$

where $\gamma > 0$ is a friction parameter. With these choices, the evolution equation (2.9) becomes

$$\dot{q} = L_{qp} \partial_p E = \frac{p}{m},$$

$$\dot{p} = L_{pq} \partial_q E - \frac{1}{2} M_{pe} \partial_e S = -\partial_q V(q) - \gamma \frac{p}{m},$$

$$\dot{e} = -\frac{1}{2} M_{ee} \partial_e S = \gamma \frac{p^2}{m^2}.$$

Note that the degeneracy conditions $L dS = 0$ and $M dE = 0$ are satisfied by construction.

Remark 2.8. The second statement in (2.7) is equivalent to the notion that $dE$ is a distinguished function for the bracket generated by $\Psi^*$ introduced on page 6624 of [1].

2.3 Pre GENERIC

We will see below that a less constrained version of GENERIC, which we will call pre GENERIC, emerges naturally for a broad class of evolution equations from their microscopic origins. For such equations, we do not consider a flow coming from some conserved energy $E$, but replace $L(z) dE(z)$ by a general flow $W(z)$. The constraint in (c) of Definition 2.5 can then be replaced by the less constraining orthogonality condition $\langle W(z), dS(z) \rangle = 0$. 
Definition 2.9. A pre-GENERIC equation with respect to a functional $S \in C^1(Z)$, dissipation potentials $(\Psi, \Psi^*)$ and vector field $W : Z \to TZ$ is given by

$$\Psi(z, \dot{z} - W(z)) + \Psi^* \left( z, -\frac{1}{2}dS(z) \right) + \frac{1}{2} (\dot{z}, dS(z)) = 0$$

(2.10)

where

(a) $(\Psi, \Psi^*)$ are symmetric;
(b) $S$ and $W$ satisfy the non-degeneracy condition $\langle W(z), dS(z) \rangle = 0$ for all $z \in Z$.

The computation in (2.8) does not change under the generalization from GENERIC to pre-GENERIC. Thus, we conclude that also in the pre-GENERIC setting the functional $S$ is decreasing under the flow.

Example 2.10. The gradient-flow equation (2.4) can also be converted into a pre-GENERIC system by adding a flow field $W$,

$$\dot{z} = W(z) - \frac{1}{2} MdS(z),$$

which should satisfy the more limited degeneracy condition $\langle W, dS \rangle = 0$.

A pre-GENERIC system is GENERIC if it satisfies the following additional conditions:

1. $W(z) = L(z)dE(z)$ for some energy functional $E \in C^1(Z)$ and some family of antisymmetric operators $L$ that satisfy the Jacobi identity;
2. $L(z)dS(z) = 0$.

3 Microscopic origins of variational evolution and fluctuation symmetries

The discussion of the papers [3, 13] shows that the gradient flow and GENERIC structures can be derived from the microscopic origins of the macroscopic evolution equation. We will follow a similar route, and show that the construction of dissipation potentials $(\Psi, \Psi^*)$ from the Hamiltonian or the Lagrangian of the path-space large deviations as in [13] for the gradient flow setting, can be extended to the setting of (generalized) GENERIC. We will recover the result of [3] as a special case.

We start by introducing the path-space large deviations.

3.1 Microscopic origins of macroscopic equations and large deviations of processes

Consider a sequence of Markov processes $X_n$ with infinitesimal generators $(A_n, D(A_n))$. The macroscopic limiting equation of the processes $X_n$ can be derived via the convergence of generators. In particular, if the limiting system is deterministic, we can usually find a set of functions $D(A) \subseteq C_c(Z)$ and functions $f_n \in D(A_n)$ that converge boundedly and uniformly on compacts(buc) to $f$ and such that $A_nf_n$ converges buc to $Af$. In addition, this limiting operator is of the form $Af(z) = \langle F(z), df(z) \rangle$ where $F$ is a vector field on $Z$. 

9
Path-space large deviations give us the fluctuations around $\dot{z} = F(z)$ of the type
\[
\mathbb{P}\{X_n(t)\}_{t \geq 0} \approx \{\gamma(t)\}_{t \geq 0} \sim \exp \left\{ -r_n \left( I_0(\gamma(0)) + \int_0^\infty \mathcal{L}(\gamma(s), \dot{\gamma}(s))ds \right) \right\}. \tag{3.1}
\]
Here $r_n$ denotes the rate of exponential decay, which could have the interpretation simply $r_n = n$, but could also be in terms of a volume depending on $n$, e.g. $r_n = n^3$. Feng and Kurtz describe an algorithm \cite{4} to find the Lagrangian $\mathcal{L} : TZ \to [0, \infty]$ as follows. One defines pre-limit Hamiltonians $H_n f = r_n^{-1} e^{r_n f} A_n e^{r_n f}$ with domain $\mathcal{D}(H_n) = \{ f | e^{r_n f} \in \mathcal{D}(A_n) \}$ and shows that there is a set of functions $\mathcal{D}(H)$ such that for $f \in \mathcal{D}(H)$ one can find $f_n \in \mathcal{D}(H_n)$ that converge buc to $f$ and such that $H_n f$ converge buc to $H f$. As above, we usually find that $H$ is of the form $H f(z) = \mathcal{H}(z, df(z))$, where $\mathcal{H}$ is a map from $T^*Z$ to $\mathbb{R}$. The Lagrangian $\mathcal{L}$ is then found by taking the Legendre transform of $\mathcal{H}$.

The associated macroscopic equation is given by the zero-cost flow: $\dot{z} = F(z)$, where $F(z) = v$ if and only if $\mathcal{L}(z, v) = 0$.

In the setting that the processes $X_n$ are reversible with respect to measures $\pi_n$, the Lagrangian respects this reversibility. Let $S$ be the rate function of, if it is satisfied, the large deviation principle with speed $r_n$. In that case, we obtain from (3.1) with $X_n(0)$ distributed as $\pi_n$ that
\[
S(\gamma(0)) + \int_0^t \mathcal{L}(\gamma(s), \dot{\gamma}(s))ds = S(\gamma(t)) + \int_0^t \mathcal{L}(\gamma(s), -\dot{\gamma}(s))ds
\]
for all $t \geq 0$ and trajectories $\gamma : [0, t] \to \mathbb{R}$ with finite cost. Infinitesimally, we find the fluctuation symmetry
\[
\mathcal{L}(z, v) - \mathcal{L}(z, -v) = \langle v, d\mathcal{S}(z) \rangle, \quad \forall (z, v) \in TZ. \tag{3.2}
\]

This relation can be taken as a starting point to show that $\dot{z} = F(z)$ is a gradient flow for $S$. This has been carried out in \cite{13}. We will recap a part of the arguments of \cite{13} and then proceed to show how GENERIC can be obtained from a generalized fluctuation symmetry.

The arguments in the next sections are purely based on the information on the mesoscopic scale, i.e. on $\mathcal{L}, H$ and $S$. The microscopic information in the processes $X_n$ will not concern us any longer.

### 3.2 Gradient flows and fluctuation symmetry

In this section, we give a recap of a part of the arguments in \cite{13}. In the following, we will always consider Hamiltonians $\mathcal{H} : T^*Z \to \mathbb{R}$ and $\mathcal{L} : TZ \to [0, \infty]$ that are convex and Legendre duals of each other. Additionally, we assume $\mathcal{L} \geq 0$ and that for each $z \in Z$ there is some $v \in T_zZ$ such that $\mathcal{L}(z, v) = 0$. Equivalently, we assume that $\mathcal{H}(z, 0) = 0$ for all $z \in Z$.

We first turn to obtaining an equivalent characterization of the fluctuation symmetry of $\mathcal{L}$ in terms of the Hamiltonian $\mathcal{H}$.

**Definition 3.1.** Let $S \in C^1(Z)$ and let $\mathcal{H} : T^*Z \to \mathbb{R}$. We say that the Hamiltonian $\mathcal{H}$ is reversible with respect to $S$ if $\mathcal{H}(z, \xi) = \mathcal{H}(z, dS(z) - \xi)$ for all $(z, \xi) \in T^*Z$.

**Proposition 3.2.** Let $V$ be a co-vector field, i.e. $V(z) \in T^*_zZ$. The following are equivalent.
(a) $\mathcal{L}$ satisfies the fluctuation symmetry
\[
\mathcal{L}(z, v) - \mathcal{L}(z, -v) = \langle V(z), v \rangle, \quad \forall (z, v) \in TZ,
\]
(b) We have
\[ H(z, \xi) = H(z, V(z) - \xi), \quad \forall (z, \xi) \in T^*Z. \]

In particular, we find that \( H \) is reversible with respect to \( S \) if and only if \( L \) satisfies the fluctuation symmetry with \( V = dS \) as in (3.2).

**Proof.** The result follows from Proposition 2.1 in [13].

**Proposition 3.3.** Let \( S \in C^1(Z) \) and let \( H : T^*Z \to \mathbb{R} \). Suppose \( H \) is reversible with respect to \( S \). Then \( S \) is a Lyapunov function for the zero-cost flow.

**Proof.** Let \( \dot{z}(t) = F(z(t)) \), where \( F(z) \) denotes the zero-cost flow field. Then
\[ \frac{d}{dt} S(z(t)) = \langle dS(z(t)), F(z(t)) \rangle = L(z(t), F(z(t))) - L(z(t), -F(z(t))) \leq 0. \]

The zero-cost flow corresponding to the a reversible Hamiltonian: \( L(z, v) = 0 \Longleftrightarrow v = \partial_\xi H(z, 0) \), can be rewritten, using the symmetry property obtained from the reversibility by
\[ v = \partial_\xi H(z, 0) = -\partial_\xi H(z, -dS(z)). \]

This is a first hint that \( H \) can be used to write the flow as a gradient flow. To extend beyond this basic property and write the flow in terms of dissipation potentials \( (\Psi, \Psi^*) \) that are symmetric and non-negative, the Hamiltonian needs to be appropriately shifted. Note that for a reversible Hamiltonian \( H \) we have
\[ \arg\min_{\xi} H(z, \xi) = \frac{1}{2} dS(z). \]

Recentering the Hamiltonian around this minimal value, we obtain a dissipation potential \( \Psi^* : T^*Z \to \mathbb{R}^+ \):
\[ \Psi^*(z, \xi) = H\left(z, \xi + \frac{1}{2} dS(z)\right) - H\left(z, \frac{1}{2} dS(z)\right). \quad (3.3) \]

\( \Psi \) is obtained from \( \Psi^* \) by taking the Legendre transform. By construction, \( \Psi \) and \( \Psi^* \) are symmetric, convex and non-negative. Also, \( v = \partial_\xi H(z, 0) = \partial_\xi \Psi^* \left(z, -\frac{1}{2} dS(z)\right) \), which is a hallmark of the gradient flow structure. Note that (3.3) is equivalent to
\[ H(z, \xi) := \Psi^* \left(z, \xi - \frac{1}{2} dS(z)\right) - \Psi^* \left(z, -\frac{1}{2} dS(z)\right). \quad (3.4) \]

The arguments in [13] show that the zero-cost flow corresponding to a reversible Hamiltonian is the gradient flow for \( S \) with respect to the dissipation potentials \( (\Psi, \Psi^*) \) and vice-versa.

**Theorem 3.4** (Theorem 2.1 in [13]). Let \( (\mathcal{L}, H) \) and \( (\Psi, \Psi^*) \) be two pairs of Legendre duals where \( H \) and \( \Psi^* \) are functionals related as in (3.3) and (3.4). Let \( S \in C^1(Z) \) and let \( F \) be some vector field.

Then the following two statements are equivalent:
(a) \( H \) is a reversible Hamiltonian for \( S \) and \( F \) is the zero-cost flow for \( \mathcal{L} \).
(b) \( \Psi^* \) is symmetric and \( F \) is the gradient-flow for \( S \) with respect to \( (\Psi, \Psi^*) \).

We give the proof for completeness and later reference.
Proof. We first prove that (a) implies (b). The zero-cost flow satisfies $L(z, F(z)) = 0$, and for all $v$ we have $L(x, v) \geq 0$. Therefore

$$0 \leq L(x, v) = \sup_{\xi} \{\langle v, \xi \rangle - H(z, \xi)\}$$

$$= \sup_{\xi} \{\langle v, \xi \rangle - \Psi^*(z, \xi)\} - H\left(z, \frac{1}{2}dS(z)\right) + \frac{1}{2}\langle v, dS(z) \rangle$$

$$= \Psi(z, v) + \Psi^*\left(z, -\frac{1}{2}dS(z)\right) + \frac{1}{2}\langle v, dS(z) \rangle.$$ 

This implies that indeed $\dot{z} = F(z) = \partial_\xi H(z, 0)$ is a gradient flow for $S$ with respect to the dissipation potentials $(\Psi, \Psi^*)$.

We proceed with the proof that (b) implies (a). The reversibility of $H$ with respect to $S$ follows from the symmetry of the functional $\Psi^*$. The gradient flow $F(z)$ satisfies by (2.3a)

$$F(z) = \partial_\xi \Psi^* \left(z, -\frac{1}{2}dS(z)\right).$$

As $H$ is defined by a shift of $\Psi^*$, we find $F(z) = \partial_\xi H(z, 0)$, which is equivalent to $L(z, F(z)) = 0$. 

The goal of this paper is to establish a similar connection in the setting of (pre-) GENERIC. We will establish equivalent conditions for the Lagrangian and Hamiltonian in terms of symmetries, that for the zero-cost trajectory leads to an equation that can be suitably interpreted as pre-GENERIC.

3.3 Pre-GENERIC and fluctuation symmetry around a vector field

Corresponding to the equivalence given between the reversibility of a Hamiltonian with respect to an entropy and the time-symmetry relation (3.2) for the associated Lagrangian, we start out with a fluctuation-symmetry that will be a first step towards GENERIC.

Motivated by the examples considered in [3], we assume that our flow consists of a gradient part for some entropy $S \in C^1(Z)$, to which we add a vector field. The fluctuation symmetry will be as in (3.2), but taken around the vector field $W$:

$$L(z, W(z) + v) - L(z, W(z) - v) = \langle dS(z), v \rangle, \quad \forall(z, v) \in TZ. \quad (3.5)$$

As has been done in [13] for the first fluctuation symmetry, we establish an equivalent description in terms of the Hamiltonian. Here the important observation, detailed by the proposition below, is that the act of translating the symmetry from the origin to $W$, as in moving from (3.2) to (3.5), is equivalent to adding a corresponding term to the Hamiltonian.

Proposition 3.5. Consider $L : TZ \to [0, \infty]$ and its Legendre transform $H : T^*Z \to \mathbb{R}$. Let $W$ be a vector field and $V$ a co-vector field, i.e. $W(z) \in T_zZ$ and $V(z) \in T^*_zZ$. The following are equivalent:

(a) We have a decomposition $H = H_1 + H_2$ such that $H_1(z, \xi) = \langle \xi, W(z) \rangle$ and $H_2$ is symmetric around $\frac{1}{2}V$: $H_2(z, \xi) = H_2(z, V(z) - \xi)$.

(b) We have $L(z, W(z) + v) - L(z, W(z) - v) = \langle V(z), v \rangle$. 

12
Proof. Let \( \mathcal{L}_2 \) denote the Legendre transform of \( \mathcal{H}_2 \). Then, we have

\[
\mathcal{L}(z, W(z) + v) = \sup_{\xi} \{ \langle \xi, W(z) + v \rangle - \mathcal{H}(z, \xi) \}
\]

\[
= \sup_{\xi} \{ \langle \xi, v \rangle - \mathcal{H}_2(z, \xi) \}
\]

\[
= \mathcal{L}_2(z, v)
\]

for all \((z, v) \in TZ\). Thus, the equivalence is derived via the application of Proposition 3.2 for \( \mathcal{L}_2 \).

We will use the result above for the setting where \( V = dS \). In particular settings, e.g. for GENERIC, we will use a specific \( W \) also, e.g. \( W = LdE \).

If GENERIC is to be found via microscopic origins as in the gradient-flow setting, we expect \( S \) to be a Lyapunov function for the zero-cost flow as in (2.8). A decomposition of the type above with \( V = dS \) is not sufficient for \( S \) to be a Lyapunov function for the zero-cost flow. In the argument of Proposition 3.3, we used that \( \mathcal{L}(z, F(z)) = 0 \). If we aim to use a similar trick in this setting, we need to shift our inner-product \( \langle dS(z), dF(z) \rangle \). Let \( \dot{z}(t) = F(z(t)) \), then (writing \( z(t) = z \) for brevity)

\[
\frac{d}{dt} S(z) = \langle dS(z), F(z) \rangle
\]

\[
= \langle dS(z), W(z) \rangle + \langle dS(z), F(z) - W(z) \rangle
\]

\[
\leq \langle dS(z), W(z) \rangle. \tag{3.5}
\]

Thus, to conclude that \( S \) is a Lyapunov function for the zero-cost flow, it suffices to have \( \langle dS(z), W(z) \rangle \leq 0 \) for all \( z \). Since we are working towards microscopic origins of GENERIC, we will however assume that the vector field \( W \) and the co-vector field \( dS \) are orthogonal. This leads to the notion of pre-GENERIC in Definition 2.9, that does not take into account a conserved energy \( E \) or operators \( L \).

The following theorem generalizes Theorem 3.4 to the pre-GENERIC setting, thus establishing that pre-GENERIC is the relevant notion when considering the generalized fluctuation symmetry (3.5).

**Theorem 3.6.** Let \( (\Psi, \Psi^*) \) be a pair of Legendre duals. Let \( S \in C^1(Z) \) and let \( F \) be a vector field.

Let \( \mathcal{H}_1(z, \xi) = \langle W(z), \xi \rangle \) for some vector field \( W \) which is orthogonal to \( dS \). Let \( \mathcal{H}_2 \) be related to \( \Psi^* \) as in (3.3) and (3.4). Finally, let \( \mathcal{H} := \mathcal{H}_1 + \mathcal{H}_2 \) and let \( \mathcal{L} \) be the Legendre dual of \( \mathcal{H} \).

The following are equivalent:

(a) \( \mathcal{H}_2 \) is reversible for \( S \) and \( F \) is the zero-cost flow for \( \mathcal{L} \).

(b) \( \Psi^* \) is symmetric and \( F \) is pre-GENERIC for \( S \) with dissipation potentials \( (\Psi, \Psi^*) \) and vector field \( W \).

If (a) or (b) is satisfied, the functional \( S \) is a Lyapunov function for \( F \).

As a corollary, we obtain conditions on the Hamiltonians that are equivalent to GENERIC.
Corollary 3.7. Consider the setting of Theorem 3.6. Let $E \in C^1(Z)$ and let $L = \{L(z)\}_{z \in Z}$ be a family of operators $L(z) : T_z^*Z \to T_zZ$ that are ‘anti-symmetric’ and satisfy the Jacobi identity (cf. Definition 2.5b). Finally suppose that $W(z) = L(z)dE(z)$.

Then the following are equivalent:

(a) $H_2$ is reversible for $S$, $LdS = 0$, and the map $\alpha \mapsto H_2(z, \xi + \alpha dE(z))$ is constant; $F$ is the zero-cost flow for $L$;

(b) $\Psi^*$ is symmetric and $F$ is GENERIC for $S, E$, dissipation potentials $(\Psi, \Psi^*)$ and map $L$.

If (a) or (b) is satisfied, then $S$ is a Lyapunov function for $F$ and the functional $E$ is constant along $F$.

Proof of Theorem 3.6. We start with the proof that (a) implies (b). Let $L_2$ be the Lagrangian corresponding to $H_2$. The zero-cost flow of $L_2$ is given by $F(z) - W(z)$. As $H_2$ is reversible for $S$, we find by Theorem 3.4 (a) to (b) that $F(z) - W(z)$ is the gradient flow for $S$ with respect to $(\Psi, \Psi^*)$:

$$
\Psi(z, F(z) - W(z)) + \Psi^* \left( z, -\frac{1}{2}dS(z) \right) + \frac{1}{2} (F(z) - W(z), dS(z)) = 0.
$$

Using that $\langle W(z), dS(z) \rangle = 0$, we find that the zero-cost flow $F$ of $L$ is pre-GENERIC for $S$, dissipation potentials $(\Psi, \Psi^*)$ and vector field $W$.

We proceed with the proof that (b) implies (a). By Using that $F$ is pre-GENERIC for $S$ with dissipation potentials $(\Psi, \Psi^*)$ and vector field $W$, we find by Theorem 3.4 (b) to (a), and the property that $\langle W(z), dS(z) \rangle = 0$ that $F - W$ is the zero-cost flow for $L_2$, the Lagrangian associated to $H_2$. Thus, we have $F(z) - W(z) = \partial_\xi H_2(z, 0)$. On the other hand

$$
\partial_\xi H(z, 0) = \partial_\xi H_1(z, 0) + \partial_\xi H_2(z, 0) = W(z) + F(z) - W(z) = F(z),
$$

which implies that the pre-GENERIC flow is the zero-cost flow for $L$.

The final statement follows from (3.6) and the property that $\langle W(z), dS(z) \rangle = 0$. \qed

Proof of Corollary 3.7. Both implications are immediate. \qed

3.4 Returning to stochastic processes: Pre-GENERIC from large deviations

In Section 3.1 we explained how Hamiltonians and Lagrangians arise in the context of large-deviation principles of stochastic processes. In Sections 3.2 and 3.3 we temporarily disregarded the stochastic origin of these functionals, and focused instead on consequences of symmetry properties of these functionals for the deterministic evolutions that they generate.

We now return to the stochastic context and consider the consequences of Sections 3.2 and 3.3 for stochastic processes.

The example that we give here is in $\mathbb{R}^d$, but we believe the general structure to be much more widely valid. The infinite-dimensional examples in Section 5 illustrate this.

Let $Z = \mathbb{R}^d$, and consider a sequence of stochastic processes $Z_n$ in $\mathbb{R}^d$, with generator $A_n$. To mirror the setup of the theorems above, we assume that $A_n$ has a decomposition $A_n = A_{1,n} + A_{2,n}$ on a common core, and that the ‘sub-generators’ $A_{1,n}$ and $A_{2,n}$ independently generate stochastic processes $X_n$ and $Y_n$.

In addition, we assume that
(a) $A_n$, $A_{1,n}$, and $A_{2,n}$ share a common unique invariant measure $\pi_n$, and that $\pi_n$ satisfies a large deviation principle on $\mathbb{R}^d$ with speed $n$ and with rate function $S \in C^1(\mathbb{R}^d)$;
(b) $\pi_n$ is reversible for $A_{2,n}$;
(c) Each of the processes $X_n$, $Y_n$, and $Z_n$ satisfy a large-deviation principle on $D_{\mathbb{R}^d}(0,\infty)$ with speed $n$ and with rate function $I$ of Lagrangian form. The Lagrangians $L$, $L_1$, and $L_2$ correspond to Hamiltonians $H$, $H_1$, and $H_2$, satisfying $H = H_1 + H_2$ and
\[ Hf(z) = \mathcal{H}(z, df(z)), \quad H_1f(z) = \mathcal{H}_1(z, df(z)), \quad \text{and} \quad H_2f(z) = \mathcal{H}_2(z, df(z)); \]
(d) $\mathcal{H}_1(z, \xi) = \langle W(z), \xi \rangle$ for some vector field $W \in C(\mathbb{R}^d; \mathbb{R}^d)$ and $H_2$ is continuous and convex in the second coordinate;

**Proposition 3.8.** Consider the setup above. Then the limiting dynamics are pre-GENERIC for $S$ with vector field $W$ and dissipation potentials $(\Psi, \Psi^*)$ defined in terms of $H_2$ as in Theorem 3.6.

**Proof of Proposition 3.8.** We prove the proposition by showing that condition $(m)$ of Theorem 3.6 are satisfied with $F$ the limiting dynamics of the processes $Z_n$. Since $Y_n$ is a reversible stochastic process (assumption $(b)$), the reversibility of $H_2$ for $S$ is given by [12, Prop. 3.7].

We now show that $H_1(z, dS(z)) = \langle W(z), dS(z) \rangle$ vanishes. For a function $u_0 : \mathbb{R}^d \to \mathbb{R}$, define for $t > 0$ the value function
\[ u(t, z) := \inf_{\gamma \in \mathcal{AC}} \left\{ u_0(\gamma(0)) + \int_0^t \mathcal{L}_1(\gamma(s), \dot{\gamma}(s)) \, ds \right\}. \]

By the contraction principle we find that if $u_0$ is the rate function of the law of the process $X_n$ at time zero, then $z \mapsto u(t, z)$ is the rate function of the law of $X_n$ at time $t$. Since $S$ is the rate function of the stationary measures, we find that if $u_0 = S$ then $u(t, \cdot) = S(\cdot)$; therefore, for all $t > 0$,
\[ S(z) = \inf_{\gamma \in \mathcal{AC}} \left\{ S(\gamma(0)) + \int_0^t \mathcal{L}_1(\gamma(s), \dot{\gamma}(s)) \, ds \right\}. \]

Now $\mathcal{L}_1(z, v)$ equals 0 if $v = W(z)$, and $+\infty$ otherwise. Therefore this identity reduces to
\[ S(\gamma(t)) = S(\gamma(0)), \]
where $t > 0$ is arbitrary and $\gamma : [0, t] \to \mathbb{R}^d$ is any curve such that $\dot{\gamma}(s) = W(\gamma(s))$ for all $0 \leq s \leq t$. By taking the limit $t \to 0$ we find that $\langle dS(z), W(z) \rangle = 0$ for all $z \in \mathbb{R}^d$.

Therefore the conditions of Theorem 3.6 are satisfied, and we find that the zero-cost flow is pre-GENERIC.

\[ \square \]

4 From pre-GENERIC to GENERIC

In [5] it is shown shown that the underdamped Fokker-Planck equation can be turned into GENERIC by adding an appropriate energy variable. The underdamped Fokker-Planck equation has a pre-GENERIC structure, as we will show in Section 5 below. It turns out that
a pre GENERIC system can be systematically extended into the GENERIC structure by the addition of an auxiliary variable which takes the role of the energy \( E \), as we show below in Section 4.1.

This extension is not unique, and in certain contexts other extensions make more physical sense. This happens in the setting of the underdamped Fokker-Planck equation in which there is a physical energy that is not conserved under the dynamics. In Section 4.2, we will consider a change of coordinates that turns our GENERIC representation of the Fokker-Planck equation in a physically more sensible one that is equivalent to the one in \([3]\).

### 4.1 Adding an energy variable for the macroscopic dynamics

The structure will be as follows. Consider a pre GENERIC system with components \( S \in C^1(Z) \), \( \Psi \), and \( W \), and equation \( \dot z = W(z) + \partial_\xi \Psi^*(z, -\frac{1}{2}\partial S(z)) \). We now add an artificial variable \( a \in \mathbb{R} \), to form a pair \((z, a) \in \hat{Z} := Z \times \mathbb{R}\). (The tangent and cotangent spaces are \( T_{z,a}Z = T_zZ \times \mathbb{R} \) and \( T^*_{z,a}Z = T^*_zZ \times \mathbb{R} \).) Since a pre GENERIC system has no concept of energy, we need to introduce one; the energy will be the mapping \( \hat{E}(z, a) := a \), and since the energy should be conserved the evolution equation for \( a \) will be \( \dot{a} = 0 \). Theorem 4.1 below formulates the resulting set of equations,

\[
\begin{align*}
\dot{z} &= W(z) + \partial_\xi \Psi^*(z, -\frac{1}{2}\partial S(z)), \\
\dot{a} &= 0,
\end{align*}
\]

as a full GENERIC system.

**Theorem 4.1.** Let \( S \in C^1(Z) \). Let \( W \) be some vector field and let \( \mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2 \) with \( \mathcal{H}_i : T^*Z \to \mathbb{R} \) and \( \Psi^* : T^*Z \to \mathbb{R}^+ \) be functionals with \( \mathcal{H}_1(z, \xi) = \langle W(z), \xi \rangle \). Let \( F \) be a vector field.

Define \( \hat{F} = (F, 0) \), \( \hat{S}(z, a) = S(z) \), \( \hat{E}(z, a) = a \),

\[
\hat{\Psi}^* \left( \begin{pmatrix} z \\ a \end{pmatrix}, \begin{pmatrix} \xi \\ r \end{pmatrix} \right) = \Psi^*(z, \xi), \quad \hat{\mathcal{H}}_i \left( \begin{pmatrix} z \\ a \end{pmatrix}, \begin{pmatrix} \xi \\ r \end{pmatrix} \right) = \mathcal{H}_i(z, \xi), \quad i \in \{1, 2\},
\]

and

\[
\hat{L}(z, a) \begin{pmatrix} \xi \\ r \end{pmatrix} = \begin{pmatrix} rW(z) \\ -\langle W(z), \xi \rangle \end{pmatrix} = \begin{pmatrix} 0 & W(z) \\ -W(z) & 0 \end{pmatrix} \begin{pmatrix} \xi \\ r \end{pmatrix}.
\]

The maps \( \hat{\Psi}^* \) and \( \hat{\mathcal{H}}_2 \) are related to each other as in \([3.3]\) and \([3.4]\). Furthermore the map \( \hat{L} \) is anti-symmetric and satisfies the Jacobi identity. If \( \langle W(z), \partial S(z) \rangle = 0 \), then \( \hat{L}d\hat{S} = 0 \).

Finally, we have the following two results:

(a) If Theorem 3.6 (a) holds for \( \mathcal{H}, S \) and \( F \), then Corollary 3.7 (a) holds for \( \hat{\mathcal{H}}, \hat{S}, \hat{E}, \hat{L} \) and \( \hat{F} \).

(b) If Theorem 3.6 (b) holds for \( \Psi^*, S \) and \( F \), then Corollary 3.7 (b) holds for \( \hat{\Psi}^*, \hat{S}, \hat{E}, \hat{L} \) and \( \hat{F} \).
Example 4.2. Performing this extension on the pre GENERIC system of Example 2.10, we find

\[ \hat{E}(z, a) := a, \quad \hat{L} := \hat{L}(z, a) = \begin{pmatrix} 0 & W \\ -W & 0 \end{pmatrix}, \]

\[ \hat{S}(z, a) := S(z), \quad \hat{M} := \hat{M}(z, a) = \begin{pmatrix} M & 0 \\ 0 & 0 \end{pmatrix}, \]

and therefore

\[ \partial_t \begin{pmatrix} z \\ a \end{pmatrix} = \begin{pmatrix} 0 & W \\ -W & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} M & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -\frac{1}{2}dS \end{pmatrix}, \]

or

\[ \dot{z} = W(z) - \frac{1}{2}MdS(z), \]
\[ \dot{a} = 0. \]

The degeneracy condition \( \hat{M}d\hat{E} = 0 \) is satisfied by construction, and since \( \hat{L}d\hat{S} = (0, -(W, dS)) \), the degeneracy condition \( \hat{L}d\hat{S} = 0 \) is satisfied by the assumption that \( (W, dS) = 0 \).

As this example illustrates, the extension with an energy variable in this way is essentially trivial: no physics is added or assumed, and the only mathematical change is a change of form. This insight raises an interesting question: Does GENERIC have any significant structure in addition to the properties of pre GENERIC? If pre GENERIC can be ‘trivially’ converted into GENERIC, is there actually any difference between the two? Whatever the answer to this question, Theorem 4.1 provides a second, independent pointer to pre GENERIC as a relevant variational structure, in addition to the appearance of pre GENERIC in large-deviation principles (see Proposition 3.8).

Proof of Theorem 4.1. The relationship between \( \hat{\Psi}^* \) and \( \hat{H}_2 \) is immediate. The map \( \hat{L} \) is anti-symmetric since

\[ \left\langle \begin{pmatrix} \xi_1 \\ r_1 \end{pmatrix}, \hat{L}(z, a) \begin{pmatrix} \xi_2 \\ r_2 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} \xi_1 \\ r_1 \end{pmatrix}, \begin{pmatrix} r_2W(z) \\ -\langle W(z), \xi_2 \rangle \end{pmatrix} \right\rangle = r_2\langle W(z), \xi_1 \rangle - r_1\langle W(z), \xi_2 \rangle = -\left\langle \begin{pmatrix} \xi_2 \\ r_2 \end{pmatrix}, \hat{L}(z, a) \begin{pmatrix} \xi_1 \\ r_1 \end{pmatrix} \right\rangle. \]

The tedious proof that \( \hat{L} \) satisfies the Jacobi identity is postponed to the appendix and follows from Proposition A.1. We verify \( \hat{L}d\hat{S} = 0 \) by a simple calculation:

\[ \hat{L}(z, a)(d\hat{S}(z, a)) = \hat{L}(z, a) \begin{pmatrix} dS(z) \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \langle W(z), dS(z) \rangle \end{pmatrix} = 0. \]

We proceed with the proof of (a). First, we find

\[ \hat{L}(z, a)d\hat{E}(z, a) = \begin{pmatrix} 1 \cdot W(z) \\ -\langle W(z), 0 \rangle \end{pmatrix} = \begin{pmatrix} W(z) \\ 0 \end{pmatrix}, \]
establishing that \( \hat{H}_1 \) is given by a linear paring with \( \hat{L}d\hat{E} \). Clearly \( \hat{H}_2 \) is reversible for \( \hat{S} \), due to the same property for \( H_2 \) and \( S \). As \( d\hat{E} \) only gives a contribution in the second coordinate, and \( \hat{H}_2 \) only depends on the first coordinates, we find that \( \alpha \mapsto \hat{H}_2 \left( (z, a), (\xi, r) + \alpha d\hat{E}(z, a) \right) \) is constant.

Finally, we prove (b). Note that (2.6) is satisfied for \( \hat{S}, \hat{E} \), dissipation potentials \((\hat{\Psi}, \hat{\Psi}^*)\) and operators \( \hat{L} \), due to the fact that \( \hat{L}d\hat{E} = \hat{W}, (\hat{\Psi}, \hat{\Psi}^*) \) act as \((\Psi, \Psi^*)\) only on the first variable, \( d\hat{S}(z, a) = (dS(z), 0) \) and the pre GENERIC property of \( F(z) \). By definition of \( \hat{E} \) and \( \hat{\Psi}^* \), we find

\[
\hat{\Psi}^*((z, a), (\xi, r) + \alpha d\hat{E}(z, a)) = \hat{\Psi}^*((z, a), (\xi, r + \alpha)) = \Psi^*(z, \xi),
\]

which is constant in \( \alpha \).

**Remark 4.3.** The addition of a trivial energy variable to the dynamics can also be carried out for a microscopic system, for instance in the setting of Proposition 3.8, leading to the following Corollary.

**Corollary 4.4.** Consider the setup in Proposition 3.8 and consider the dynamics \((X_n(t), 0)\) and energy \( \hat{E}(z, a) = a \).

The processes \((X_n(t), 0)\) preserve \( \hat{E} \) and the limiting dynamics \((F, 0)\) are GENERIC with energy \( \hat{E} \), entropy \( \hat{S} \), operators \( \hat{L} \) and dissipation potentials \((\hat{\Psi}, \hat{\Psi}^*)\) as in Theorem 4.1.

**Proof.** This result is immediate from Proposition 3.8 and Theorem 4.1.

**4.2 Transforming GENERIC in the setting with a natural energy**

In Section 5 below, we will see that the underdamped Fokker-Planck equation and the Andersen thermostat are pre GENERIC system with a vector field \( W \) that can be written as \( L(z)dE(z) \) for some physical energy \( E \). In this situation the choice of Theorem 4.1 is less natural; one would want to form a GENERIC structure in which \( E(z) \) plays a role as the energy.

However, in the examples in Section 5 the energy \( E \) is not conserved by the flow. At a microscopic level, this is the consequence of energy flowing in and out of a heat bath; this heat bath is implicitly present as the generator of the noise in the system. Therefore the natural conserved energy should be of the form \( E(z) + e \), where the new variable \( e \) represents the energy of the heat bath. In [3] it was shown that this way of extending the system indeed leads to a GENERIC structure for the extended system, and here we generalize the ideas to the nonlinear dissipative structures of this paper.

To explain the choices below we first discuss the running example.

**Example 4.5.** Again starting from the pre GENERIC example of Example 2.10, we now assume that \( W = LdE \) for some antisymmetric operator \( L \) satisfying the Jacobi identity and some function \( E \in C^1(Z) \). Adding a variable \( e \) that describes the energy of the heat bath, as discussed above, we calculate that

\[
0 = \partial_t (E(z) + e) = \langle dE, \dot{z} \rangle + \dot{e} = \langle dE, LdE - \frac{1}{2}Mds \rangle + \dot{e} = -\frac{1}{2} \langle dE, Mds \rangle + \dot{e}.
\]
Therefore choosing the natural evolution $\dot{e} = \frac{1}{2}(dE, MdS)$ we are led to the system

\[
\begin{align*}
\dot{z} &= LdE(z) - \frac{1}{2}MS(z), \quad (4.1a) \\
\dot{e} &= \frac{1}{2}(dE, MdS). \quad (4.1b)
\end{align*}
\]

This system can be obtained from the system of Theorem 4.1 by the transformation $\Phi: (z, a) \mapsto (z, a - E(z)) =: (z, e)$, which replaces the full energy $a$ by the heat-bath energy $a - E(z) =: e$. This structure is given by $\bar{E}(z, e) = E(z) + e$ and $\bar{S}(z, e) = S(z)$, and

\[
\partial_t \begin{pmatrix} z \\ e \end{pmatrix} = \begin{pmatrix} 0 & LdE \\ -LdE & 0 \end{pmatrix} \begin{pmatrix} dE \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} M \\ -M \end{pmatrix} \begin{pmatrix} -MdE, \cdot \\ \langle dE, MdE \rangle \end{pmatrix} \begin{pmatrix} dS \\ 0 \end{pmatrix}.
\]

A next natural step would be to replace the first term above using the calculation

\[
\begin{pmatrix} 0 & LdE \\ -LdE & 0 \end{pmatrix} \begin{pmatrix} dE \\ 1 \end{pmatrix} = \begin{pmatrix} LdE \\ 0 \end{pmatrix} = \begin{pmatrix} L & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} dE \\ 1 \end{pmatrix},
\]

since the final product reflects more naturally the structure of the original equation for $z$. From the pre GENERIC structure, however, we need not have $LdS = 0$, implying that a structure of this type need not satisfy the degeneracy conditions. In fact, in the examples of Section 5 we will have $LdS \neq 0$ but $L(dS - dE) = 0$. Note that this is compatible with $\langle W, dS \rangle = 0$, since by the antisymmetry of $L$,

\[
\langle W, dS \rangle = \langle LdE, dS \rangle = \langle L(dE - dS), dS \rangle = 0.
\]

To regain the degeneracy conditions we therefore redefine the new entropy to be $\bar{S}(z, e) = S(z) - E(z) - e$. The resulting system,

\[
\partial_t \begin{pmatrix} z \\ e \end{pmatrix} = \begin{pmatrix} L & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} dE \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} M \\ -M \end{pmatrix} \begin{pmatrix} -MdE, \cdot \\ \langle dE, MdE \rangle \end{pmatrix} \begin{pmatrix} dS - dE \\ 1 \end{pmatrix},
\]

satisfies all the requirements of the GENERIC structure, and gives (4.1) as evolution equation.

Summarizing, we first made the choice $e = a - E(z)$ for the desired additional variable. While there is a perfectly admissible GENERIC structure for the variable $(z, e)$, we also desired that the Poisson operator be of the form $\begin{pmatrix} L & 0 \\ 0 & 0 \end{pmatrix}$. In the examples of Section 5 we do not have $LdS = 0$, only $L(dS - dE) = 0$, and to make this correspond to a degeneracy condition we translated the entropy by subtracting $E$.

The next theorem describes this procedure for the nonlinear GENERIC structure of this paper, and proves the corresponding properties.
Theorem 4.6. Let $S, E \in C^1(Z)$. Suppose $F$ is pre GENERIC for $S$ with dissipation potentials $(\Psi, \Psi^*)$ and vector field $W$. Suppose that $W(z) = L(z)dE(z)$, where

(a) $L$ are anti-symmetric operators $L(z): T^*_zZ \to T_zZ$ that satisfy the Jacobi identity,
(b) $L(z)(dS(z) - dE(z)) = 0$.

Define

$$\Theta^+(\begin{pmatrix} z \\ e \end{pmatrix}, \begin{pmatrix} \xi \\ r \end{pmatrix}) = \Psi^*(z, \xi - rdE(z)),$$

let $\Theta$ be defined as the Legendre transform of $\Theta^+$, set $\Theta(z, e) = S(z) - (E(z) + e)$, $\Theta(z, e) = E(z) + e$, and

$$L(z, e)\begin{pmatrix} \xi \\ r \end{pmatrix} = \begin{pmatrix} L(z)\xi \\ 0 \end{pmatrix}.$$

Then the flow $\overline{F}(z, e) = (F(z), -(dE(z), F(z)))$ is GENERIC for $\overline{S}, \overline{E}$ with dissipation potentials $(\Theta, \Theta^*)$ and operators $L$.

Remark 4.7. The transition from ‘unbarred’ to ‘barred’ objects in Theorem 4.6 can be extended to obtain a continuous scale of interpolating GENERIC systems. Let $\alpha \in [0, 1]$ and define $E_\alpha(z, e) = e + E(z), S_\alpha(z, e) = S(z) - \alpha(e + E(z)), \Psi^*_\alpha(z, e, (\xi, r)) = \Psi^*(z, \xi - rdE(z))$, and

$$L_\alpha(z, e) = \begin{pmatrix} \alpha L(z) & (1 - \alpha)W(z) \\ -(1 - \alpha)W(z) & 0 \end{pmatrix}.$$

Then $\overline{F}(z, e)$ is GENERIC for $S_\alpha, E_\alpha$ with dissipation potentials $(\Psi_\alpha, \Psi^*_\alpha)$ and operators $L_\alpha$.

Proof of Theorem 4.6. The proof consists of two steps. We start with the GENERIC representation of Theorem 4.6 after which we make a coordinate change in the energy variable. This gives us GENERIC in a new set of variables for a changed energy and a changed dissipation potential. Then we make a second step, simultaneously changing the entropy functional and the anti-symmetric operator.

Step 1. Denote $\tilde{Z} = Z \times \mathbb{R}$. Let $\Phi: \tilde{Z} \to \tilde{Z}$ be defined as $\Phi(z, e) = (z, e - E(z))$. Then $D\Phi: T\tilde{Z} \to T\tilde{Z}$ and $(D\Phi)^* : T^*\tilde{Z} \to T^*\tilde{Z}$ and

$$D\Phi \begin{pmatrix} (z, e), \begin{pmatrix} v \\ v_e \end{pmatrix} \end{pmatrix} = \begin{pmatrix} (z, e - E(z)), \begin{pmatrix} v \\ v_e - (dE(z), v) \end{pmatrix} \end{pmatrix},$$

$$(D\Phi)^* \begin{pmatrix} (z, e), \begin{pmatrix} \xi \\ r \end{pmatrix} \end{pmatrix} = \begin{pmatrix} (z, e + E(z)), \begin{pmatrix} \xi - dE(z) \\ r \end{pmatrix} \end{pmatrix}. \quad (4.2)$$

We obtain dynamics on $\tilde{Z}$ by pushing forward the GENERIC dynamics on $\tilde{Z}$ via the map $\Phi$. Thus, all operators change under the action of $\Phi$, and we obtain that the flow

$$\partial_t \begin{pmatrix} z(t) \\ e(t) \end{pmatrix} = \begin{pmatrix} F(z(t)) \\ -(dE(z(t), F(z(t))) \end{pmatrix} \quad (4.3)$$

is a GENERIC system in new coordinates with functionals

$$\tilde{E}(z, e) = \tilde{E} \circ \Phi^{-1}(z, e) = e + E(z),$$

$$\tilde{S}(z, e) = \tilde{S} \circ \Phi^{-1}(z, e) = S(z),$$

$$\tilde{L} \begin{pmatrix} (z, e), \begin{pmatrix} \xi \\ r \end{pmatrix} \end{pmatrix} = D\Phi \circ \tilde{L} \circ (D\Phi)^* \begin{pmatrix} (z, e), \begin{pmatrix} \xi \\ r \end{pmatrix} \end{pmatrix} = \begin{pmatrix} (z, e), \begin{pmatrix} 0 & W(z) \\ -W(z) & 0 \end{pmatrix} \begin{pmatrix} \xi \\ r \end{pmatrix} \end{pmatrix}.$$
and

\[ \tilde{\Psi}^* \left( (z, e), \left( \xi \over r \right) \right) = \tilde{\Psi}^* \left( (D\Phi)^* \left( (z, e), \left( \xi \over r \right) \right) \right) \]

\[ = \tilde{\Psi}^* \left( (z, e + E(z)), \left( \xi - r dE(z) \over r \right) \right) = \Psi^* (z, \xi - rdE(z)). \]

*Step 2.* Set \( \mathcal{E} = \tilde{E}, \mathcal{L}^* := \tilde{\Psi}^* \) and \( \mathcal{L} = \tilde{\Psi} \). Next, we change both the entropy and the operator \( L \) to obtain the desired GENERIC representation. Set \( \mathcal{S} = \tilde{S} - \tilde{E} \) and

\[ \mathcal{L} = \left( \begin{array}{ccc} L & 0 \\ 0 & 0 \end{array} \right). \]

We prove that (4.3) is also GENERIC for \((\mathcal{S}, \mathcal{E}, \mathcal{L}^*, \mathcal{L})\).

First of all, note that \( \mathcal{L} \) is anti-symmetric and satisfies the Jacobi identity because \( L \) does. Thus, we need to show that the equation in (2.6) for the GENERIC representation in step 1, implies the same equation for our new representation. In addition, we need to check the non-degeneracy condition, (2.7). We start with the first part of the latter. We have

\[ \mathcal{L} d\mathcal{S} = \mathcal{L} (d\tilde{S} - d\tilde{E}) = \left( \begin{array}{ccc} L & 0 \\ 0 & 0 \end{array} \right) \left( dS(z) - dE(z) \right) = 0 \]

by assumption (b). For the second part of (2.7) note that \( \tilde{\Psi}^* = \tilde{\Psi}^* \) is constant by addition of \( \tilde{E} \) in the momentum variable by the GENERIC property of the system obtained in step 1.

We proceed with the verification of (2.6) for our new representation by comparing the contributions of the old and new representation. In particular, we prove that

\[ \mathcal{L} d\mathcal{S} = \mathcal{L} (d\tilde{S} - d\tilde{E}) = \left( \begin{array}{ccc} L & 0 \\ 0 & 0 \end{array} \right) \left( dS(z) - dE(z) \right) = 0 \]

1/2 \( \langle \dot{\xi}, d\tilde{S}(z) \rangle = 1/2 \langle \dot{\xi}, d\tilde{S}(z) \rangle \)

For (4.4) note that \( \bar{\Psi} = \tilde{\Psi}, \tilde{E} = \tilde{E} \) and \( \mathcal{L} d\mathcal{E} = \mathcal{L} \tilde{E} \tilde{E}. \) (4.3) follows because \( \bar{\Psi} = \tilde{\Psi}^* \) and the property of \( \alpha \rightarrow \tilde{\Psi}^* ((z, e), (\xi, r) + \alpha d\tilde{E}(z, e)) \) being constant. The final equality (4.6) follows from the fact that \( \tilde{E} \) is constant under the flow by the GENERIC property for \((\mathcal{S}, \mathcal{E}, \tilde{\Psi}, \tilde{\Psi}^*). \)

5 Two related examples: the Andersen thermostat and the Vlasov-Fokker-Planck equation

5.1 The macroscopic dynamics and pre GENERIC

In this final section, we work out the main idea’s in this paper for the the Andersen thermostat and the Vlasov-Fokker-Planck equation. In both cases, the evolution equation that we are considering is given by \( d_t \rho = \mathbf{F}(\rho) \) with

\[ \mathbf{F}(\rho) = -\text{div}_q \left( \rho \over m \right) + \text{div}_p (\rho \nabla_q V + \rho (\nabla_q \Phi * \rho)) + G^*(\rho). \]
Here $\rho \in \mathcal{P}(\mathbb{R}^{2d})$ where $\mathbb{R}^{2d}$ represents the space of pairs $(q_1, \ldots, q_d, p_1, \ldots, p_d)$ of position and momentum in $\mathbb{R}^d$. We implicitly think of this vector field being defined on the tangent bundle $T\mathcal{P}(\mathbb{R}^{2d})$ with $T_{\rho}\mathcal{P}(\mathbb{R}^{2d}) = H^{-1}(\rho)$. The subscripts $p, q$ in $\text{div}_p, \text{div}_q, \Delta_p$ mean that the operators act only on the corresponding variable. $V, \Phi : \mathbb{R}^d \to \mathbb{R}^d$ are two potentials that grow sufficiently fast at infinity and we assume that $\Phi$ is symmetric. The convolution is given by

$$(\Phi * \rho)(q) := \int \Phi(q - q')\rho(dq', dp')$$

and is a function of $q$ only. The dynamics $\mathcal{G}$ is given by the adjoint of an operator $\mathcal{G}$ that can take various forms. Two main examples are given by:

**Andersen Thermostat** redistribution of the momentum via a jump-processes with a Maxwellian distribution

$$\mathcal{G} f(q, p) := \gamma \int \left[ f(q, p') - f(q, p) \right] \frac{1}{\sqrt{2\pi m}} e^{-\frac{p^2}{2m}} \, dp,$$  \hspace{1cm} \text{(5.2)}$$

**Vlasov-Fokker-Planck** a diffusion process modeling stochastic forcing with friction

$$\mathcal{G} f(q, p) := -\gamma \left\langle \frac{p}{m} \nabla_p f(q, p) \right\rangle + \gamma \Delta_p f(q, p).$$ \hspace{1cm} \text{(5.3)}$$

In both cases $m > 0$ models the mass of the particles, whereas $\gamma$ models the intensity of the interactions with a background solvent. In general, we can consider more general $\mathcal{G}$, e.g. certain types of Lévy processes. We first work out these two examples and come back to general conditions under which the analysis can be carried out in Condition 5.4 below.

At this point, it is important to note that the operator $\mathcal{G}$ acts on the momentum variable only and is independent of the value of $q$, i.e. if $f$ depends on $p$ only, then $\mathcal{G} f(q, p)$ depends on $p$ only. This assumption is natural from a physical point of view as $\mathcal{G}$ models the interaction between our particles and a background solvent.

In addition, the operator $\mathcal{G}$ interpreted as acting on $p$ only is reversible with respect to a centered Gaussian measure

$$\rho_m(dp) = \frac{1}{\sqrt{2\pi m}} e^{-\frac{p^2}{2m}} dp,$$

with $m > 0$.

We introduce some additional structure to study the macroscopic equations in relation to pre GENERIC and GENERIC. First of all, define $S, E : \mathcal{P}(\mathbb{R}^{2d}) \to \mathbb{R}$ by

$$E(\rho) := \int \frac{p^2}{2m} + V(q) + \frac{1}{2} (\Phi * \rho)(q) \, \rho(dq, dp),$$ \hspace{1cm} \text{(5.4)}$$

$$S(\rho) := \int \log \rho(q, p) \, \rho(dq, dp) + E(\rho).$$ \hspace{1cm} \text{(5.5)}$$

which have functional derivatives

$$dE(\rho)(q, p) = \frac{p^2}{2m} + V(q) + (\Phi * \rho)(q),$$ \hspace{1cm} \text{(5.6)}$$

$$dS(\rho)(q, p) = \log \rho(q, p) + 1 + dE(\rho).$$ \hspace{1cm} \text{(5.7)}$$

Next, let $J$ be the $2d$-dimensional symplectic matrix

$$J = \begin{bmatrix} 0 & -\mathbb{1}_d \\ \mathbb{1}_d & 0 \end{bmatrix},$$

22
and define $L : T^* \mathcal{P}(\mathbb{R}^{2d}) \to T^* \mathcal{P}(\mathbb{R}^{2d})$ by

$$L(\rho)\xi := \text{div}(\rho J \nabla \xi).$$

Note that $L$ inherits the symplectic structure of $J$, and is therefore a family of anti-symmetric operators satisfying the Jacobi identity.

**Theorem 5.1.** The Andersen thermostat and the Vlasov-Fokker-Planck equation are pre-Generic with entropy $S$, vector field $W(\rho) := L(\rho)dE(\rho)$ and dissipation potentials

$$\Psi^*(\rho, \xi) := \gamma \int \int [\cosh(\xi(q, p') - \xi(q, p)) - 1] \frac{1}{\sqrt{2\pi m}} e^{-\frac{(p^2 + (p')^2)}{4m}} \sqrt{\rho(q, p') \rho(q, p)} dp' dp dq,
$$

$$\Psi(\rho, \xi) := \gamma |\xi|_H^2(\rho) = \gamma \int |\nabla_p \xi(q, p)|^2 d\rho(q, p),$$

respectively.

Before giving the proof of the result, we start with some preliminary properties of the drift and the operators $L$ that are needed at various points in this section.

**Lemma 5.2.** We have

(a) \]

$$W(\rho) := L(\rho)dE(\rho) = - \text{div}_q \left( \frac{p}{m} \right) + \text{div}_p \left( \rho \nabla_q V + \rho(\nabla_q \Phi * \rho) \right),$$

(b) \]

$$L(\rho)(dS(\rho) - dE(\rho)) = 0.$$

**Proof.** Both claims follow by direct calculation:

$$L(\rho)dE(\rho) = \text{div} J \nabla dE(\rho) = \text{div}_p \left( \rho \nabla_q V(q) + (\Phi * \rho)(q) \right) - \text{div}_q \left( \frac{p}{m} \right),$$

and

$$L(\rho)(dS(\rho) - dE(\rho)) = \text{div} J \nabla (\log \rho) = \text{div}_p \left( \rho \nabla_q \log \rho \right) - \text{div}_q \left( \rho \nabla_p \log \rho \right)
= \text{div}_p \left( \nabla_q \rho \right) - \text{div}_q \left( \nabla_p \rho \right) = 0.$$

**Proof of Theorem 5.1.** We start by checking property (a) and (b) of Definition 2.9. Clearly, for both the Andersen thermostat and the Vlasov-Fokker-Planck equation the potential $\Psi^*$ is non-negative, symmetric and $\Psi^*(\rho, 0) = 0$. By the properties of the Legendre transform the same follows for $\Psi$. Property (b) follows by using the anti-symmetry of $L$ twice and then Lemma 5.2 (b) (we drop dependence on $\rho$):

$$\langle W, dS \rangle = \langle LdE, dS \rangle = \langle LdE, dS - dE \rangle = -\langle dE, L(dS - dE) \rangle = 0.$$
We start with the calculation of $\partial_\xi \Psi^*(\rho, \xi)$, which is given by the property that for any $\eta$ in the cotangent space, we have

$$\langle \eta, \partial_\xi \Psi^*(\rho, \xi) \rangle = \lim_{t \to 0} \frac{\Psi^*(\rho, \xi + t\eta) - \Psi^*(\rho, \xi)}{t}.$$ 

For the Andersen Thermostat, we have

$$\lim_{t \to 0} \frac{\Psi^*(\rho, \xi + t\eta) - \Psi^*(\rho, \xi)}{t} = \gamma \int \left[ \eta(q, p') - \eta(q, p) \right] \sinh \left( \xi(q, p') - \xi(q, p) \right) \frac{1}{\sqrt{2\pi m}} e^{-\frac{(p')^2}{2m}} \sqrt{\rho(q, p')\rho(q, p)} \ dp' \ dp \ dq.$$

Expanding the hyperbolic sine, we find

$$\langle \eta, \partial_\xi \Psi^*(\rho, \xi + \frac{1}{2}dS(\rho)) \rangle = \gamma \int \left[ \eta(q, p') - \eta(q, p) \right] \frac{1}{\sqrt{2\pi m}} e^{-\frac{(p')^2}{2m}} \rho(q, p') \ dp' \ dp \ dq.$$ 

For the Vlasov-Fokker-Planck case, we have

$$\Psi^*(\rho, \xi + t\eta) - \Psi^*(\rho, \xi) = 2t\gamma \int \nabla_p \xi \cdot \nabla_p \eta \ dp + t^2\gamma \int \nabla_p \eta \cdot \nabla_p \eta \ dp.$$

Dividing by $t$ and sending $t$ to zero, we find

$$\langle \eta, \partial_\xi \Psi^*(\rho, \xi) \rangle = 2\gamma \int \nabla_p \xi \cdot \nabla_p \eta \ dp = \langle \eta, -2 \text{div}_p (\rho \nabla_p \xi) \rangle$$

establishing that $\partial_\xi \Psi^*(\rho, \xi) = -2 \text{div}_p (\rho \nabla_p \xi)$. Filling out $-\frac{1}{2}dS(\rho)$, we obtain

$$\partial_\xi \Psi^* \left( \rho, -\frac{1}{2}dS(\rho) \right) = -2 \text{div}_p \left( \rho \nabla_p \frac{1}{2}dS(\rho) \right) = \gamma \text{div}_p \left( \rho \nabla_p \left( \log \rho(q, p) + 1 + \frac{p^2}{2m} \right) \right)$$

$$= \gamma \text{div}_p \left( \nabla_p \rho(q, p) + \rho(q, p) \frac{p}{m} \right)$$

$$= \gamma \Delta_p \rho(q, p) + \gamma \text{div}_p \left( \rho(q, p) \frac{p}{m} \right)$$

$$= \mathcal{G}^*(\rho).$$

We will show that the dynamics \ref{eq:5.8} can be obtained from a microscopic model of interacting particles which via their large deviations give rise to a pre GENERIC structure that can be extended in a natural way to GENERIC.

It should be noted that we will not consider the difficulties involved in proving path-space large deviation principles for the models considered.
5.2 Microscopic model and its description in terms of macroscopic variables

We consider \( n \) interacting particles with position and momentum \((Q_i, P_i) \in \mathbb{R}^d \times \mathbb{R}^d\). The dynamics of our particles consist of three separate parts:

(a) independent fluctuations of the momentum variables, modeling the interaction with some background solvent;
(b) a conservative part that describes the interplay between position and momentum;
(c) pair interactions between the particles that influence the momentum.

**Part (a).** The fluctuation of the momentum variable is described by a generator \(G \subseteq C_b(\mathbb{R}^d) \times C_b(\mathbb{R}^d)\), given by (5.2) for the Andersen Thermostat and (5.3) for the Vlasov-Fokker-Planck case.

**Part (b).** For the interplay between position and momentum, we consider dynamics for which each particle, independent of every other particle, is given by the generator

\[
B f(q, p) := \langle \frac{p}{m}, \nabla_q f(q, p) \rangle - \langle \nabla_q V(q), \nabla_p f(q, p) \rangle
\]

modeling the speed proportional to the momentum, and the change in momentum under the influence of the potential \(V\).

**Part (c).** Finally, we consider the interaction amongst particles. The dynamics of the \( i \)-th particle depends on the locations of the other particles, and this dependence is represented by the generator

\[
C_i f(q_1, p_1, \ldots, q_n, p_n) := -\frac{1}{2n} \sum_{j=1}^{n} \nabla \Phi(q_i - q_j) \nabla_{p_i} f(q_1, p_1, \ldots, q_n, p_n).
\]

Combining these operators, writing \(y_i = (q_i, p_i)\), the total dynamics \(\{Y_i(t)\}_{1 \leq i \leq n, t \geq 0} = \{(Q_i(t), P_i(t))\}_{1 \leq i \leq n, t \geq 0}\) has generator

\[
A_n f(y_1, \ldots, y_n) = \sum_{i=1}^{n} (G_f(y_1, \ldots, y_{i-1}, y_i, y_{i+1}, \ldots, y_n)) (y_i)
\]

\[
+ \sum_{i=1}^{n} (B f(y_1, \ldots, y_{i-1}, y_i, y_{i+1}, \ldots, y_n)) (y_i)
\]

\[
+ \sum_{i=1}^{n} C_i f(y_1, \ldots, y_n).
\]

(5.9)

Our goal is to show that the dynamics in (5.9), after taking empirical densities, converges to the solution of (5.1).

Let \(\eta_n : (\mathbb{R}^{2d})^n \to \mathcal{P}(\mathbb{R}^{2d})\) be the map

\[
\eta_n(y_1, \ldots, y_n) := \frac{1}{n} \sum_{j=1}^{n} \delta_{y_j}
\]

for the empirical measure of \(n\) particles. Since the action of the operators \(G\) and \(B_n\) on the \(i\)-th particle only depends on the particle itself, and the interaction of the \(i\)-th particle with other particle dynamics as induced by \(C_i\) can be written in terms of the empirical density,

\[
C_i f(y_1, \ldots, y_n) := -\langle \nabla \Phi * \eta_n(y_1, \ldots, y_n)(q_i), \nabla_{p_i} f(y_1, \ldots, y_n) \rangle,
\]

25
Hereafter an application of (5.10), yields for the Andersen thermostat the dynamics of \( \{ \eta_n(Y_1(t), \ldots, Y_n(t)) \}_{t \geq 0} \) is autonomous. For a suitable class of functions \( F \) on \( \mathcal{P}(\mathbb{R}^{2d}) \) the generator \( A_n \) of these dynamics satisfies

\[
A_n F(\eta_n(y_1, \ldots, y_n)) = [A_n(F \circ \eta_n)](y_1, \ldots, y_n).
\tag{5.10}
\]

To obtain an explicit representation for \( A_n \), we first introduce a convenient class of test functions on \( \mathcal{P}(\mathbb{R}^{2d}) \). We say that \( F \) is a cylinder function if there are \( f_1, \ldots, f_l \in C^2_b(\mathbb{R}^{2d}) \) and \( \phi \in C^2(\mathbb{R}^l) \) so that

\[
F(\rho) = \phi((f_1, \rho), \ldots, (f_l, \rho)) =: \phi((f, \rho))
\]

Here \( f \) denotes the vector \((f_1, \ldots, f_l)\) and \((f, \rho)\) denotes the vector of integrals \((\rho, f_1, \ldots, \rho, f_l)\).

Note that the functional derivative of the generator \( G \) given by

\[
\gamma \int \left[ F \left( \rho - \frac{1}{n} \delta(q_1, \rho_1) + \frac{1}{n} \delta(q_2, \rho_2) \right) - F(\rho) \right] \frac{1}{\sqrt{2\pi m}} e^{-\frac{x^2}{2m}} \, dp'
\]

\[
= \gamma \int \left[ F \left( \rho - \frac{1}{n} \delta(q_1, \rho_1) + \frac{1}{n} \delta(q_2, \rho_2) \right) - F(\rho) \right] \frac{1}{\sqrt{2\pi m}} e^{-\frac{x^2}{2m}} \, dp' \rho(d(q, p)),
\tag{5.12}
\]

and for the Vlasov-Fokker-Planck equation

\[
G_n F(\rho) = \sum_{i=1}^{l} \partial_i \phi((f, \rho)) \langle \mathcal{G}_f, \rho \rangle + \frac{2}{n} \sum_{i,j=1}^{l} \partial_i \partial_j \phi((f, \rho)) \langle \nabla_p f_i \nabla_p f_j, \rho \rangle,
\tag{5.13}
\]

\[
= \langle dF(\rho), \mathcal{G}^\ast(\rho) \rangle + \frac{\gamma}{n} \sum_{i,j=1}^{l} \partial_i \partial_j \phi((f, \rho)) \langle \nabla_p f_i \nabla_p f_j, \rho \rangle.
\]

Taking the limit in both cases, we find \( \lim_n G_n F(\rho) = \langle dF(\rho), \mathcal{G}^\ast(\rho) \rangle \). The operators \( B_n, C_n \) given by

\[
B_n F(\rho) = \langle \mathcal{B}(dF(\rho)), \rho \rangle,
\]

\[
C_n F(\rho) = -\int (\nabla q \Phi \ast q \rho, \nabla_p dF(\rho)) d\rho,
\]

26
are independent of \( n \) and thus have limits \( \langle dF(\rho), B^*(\rho) \rangle \) and \( \langle dF(\rho), C^*(\rho) \rangle \) respectively, where \( C^*(\rho) := \text{div}_p (\rho \nabla_q (\Phi \ast q)) \). Thus, the dynamics on the level of the empirical measure is given by the generators \( A_n = G_n + B_n + C_n \). In the limit, we find a first order generator

\[
\lim n A_n F(\rho) = \langle dF(\rho), F(\rho) \rangle = \langle dF(\rho), B^*(\rho) + C^*(\rho) + G^*(\rho) \rangle
\]
corresponding to the dual formulation of (5.1).

5.3 Stationary and path-space large deviations

A straightforward calculation using integration by parts yields that the measure

\[
\mu_n(dq_1, dp_1, \ldots, dq_n, dp_n) = Z_n^{-1} e^{-\left( \sum_{i=1}^{n} \nu(q_i) + \frac{1}{2m} \rho_1^2 + \frac{1}{m} \sum_{i<j} \Phi(q_i-q_j) \right)} \prod_{i=1}^{n} dq_i dp_i \quad (5.14)
\]
is stationary for the dynamics generated by \( A_n \) as in (5.9).

**Lemma 5.3.** Consider the stationary measures \( \mu_n \) of (5.14). The measures \( \mu_n \circ \eta_n^{-1} \) satisfy the large deviation principle on \( \mathcal{P}(\mathbb{R}^d) \) at speed \( n \) with rate function \( S \) as in (5.5).

**Proof.** The measure in (5.14) can be rewritten as

\[
\mu_n(dq_1, \ldots, dq_n) = Z_n^{-1} e^{\left( \nu(q_1) + \frac{1}{2m} \rho_1^2 + \frac{1}{m} \sum \Phi(q_1-q_j) \right)} \prod_{i=1}^{n} dq_i.
\]

Thus, taking out the dependent parts involving \( \Phi \), applying Sanov’s theorem, and then using Varadhan’s Lemma to re-insert the dependent parts, the result follows. \( \square \)

For the path-space large deviations, we follow the approach as in Section 3.1. First, we calculate

\[
H_n F(\rho) = \frac{1}{n} e^{-nF(\rho)} \left( A_n e^{nF} \right)(\rho)
\]
for cylinder functions. For the Andersen Thermostat, using (5.12), we find

\[
H_n F(\rho) = B_n F(\rho) + C_n F(\rho)
\]

\[
+ \gamma \int \int \left( \exp \left\{ \frac{F \left( \rho - \frac{1}{n} \delta(q,p) - \frac{1}{n} \delta(q,p') \right) - F(\rho)}{n^{-1}} \right\} - 1 \right) \frac{1}{\sqrt{2 \pi m}} e^{-\frac{\rho' p^2}{2m}} dp' \rho(d(q,p))
\]
which has, for good differentiable functions \( F \) with \( dF \in T^* Z \), a limit

\[
HF(\rho) = \langle dF(\rho), B^*(\rho) + C^*(\rho) \rangle
\]

\[
+ \gamma \int \int \left( \exp \left\{ \langle dF(\rho), \delta(q,p') - \delta(q,p) \rangle \right\} - 1 \right) \frac{1}{\sqrt{2 \pi m}} e^{-\frac{\rho' p^2}{2m}} dp' \rho(d(q,p))
\]

Note that this expression only depends on \( F \) via its functional derivative \( dF \) at \( \rho \): \( HF(\rho) = \mathcal{H}(\rho, dF(\rho)) \), where \( \mathcal{H} : T^* Z \to \mathbb{R} \) is defined as

\[
\mathcal{H}(\rho, \xi) = \langle \xi, B^*(\rho) + C^*(\rho) \rangle
\]

\[
+ \gamma \int \int \left[ \exp \{ \langle \xi(q,p') - \xi(q,p) \rangle \} - 1 \right] \frac{1}{\sqrt{2 \pi m}} e^{-\frac{\rho' p^2}{2m}} dp' \rho(d(q,p)).
\]
For the Vlasov-Fokker-Planck equation, we find, using (5.13), that
\[ H_n F(\rho) = \langle dF(\rho), B^*(\rho) + C^*(\rho) + G^*(\rho) \rangle + \gamma \sum_{i,j=1}^l \partial_i \phi((f, \rho)) \partial_j \phi((f, \rho)) (\nabla_p f_i \nabla_p f_j, \rho) \]
\[ + \frac{\gamma}{n} \sum_{i,j=1}^l \partial_i \partial_j \phi((f, \rho)) (\nabla_p f_i \nabla_p f_j, \rho) \]
with limiting \( HF(\rho) \) given by
\[ HF(\rho) = \langle dF(\rho), B^*(\rho) + C^*(\rho) + G^*(\rho) \rangle + \gamma \| dF(\rho) \|_{H_{\rho}}^2, \]
where
\[ \| \xi \|^2_{H_{\rho}} := \int |\nabla_p \xi(q, p)|^2 \rho(q, p). \]
Again, the Hamiltonian takes the form \( HF(\rho) = \mathcal{H}(\rho, dF(\rho)) \) with
\[ \mathcal{H}(\rho, \xi) := \langle \xi, B^*(\rho) + C^*(\rho) + G^*(\rho) \rangle + \gamma \| \xi \|^2_{H_{\rho}} = \langle \xi, F(\rho) \rangle + \gamma \| \xi \|^2_{H_{\rho}}. \]

### 5.4 Pre-generic from the large deviations

Since the large deviations of our microscopic model are the combination of a reversible dynamics \( \mathcal{G} \) and deterministic dynamics \( B + C \) that preserve \( E \), we expect that the dissipation potentials \( \Psi^* \) of Theorem 5.1 are indeed found from the Hamiltonian \( H_2 \) obtained from \( \mathcal{G} \). We calculate \( \Psi^* \) by (5.3).

For the Andersen thermostat, we find
\[
\Psi^*(\rho, \xi) = H_2 \left( \rho, \xi + \frac{1}{2} dS(\rho) \right) - H_2 \left( \rho, \frac{1}{2} dS(\rho) \right)
= \gamma \int \int [\exp \{\xi(q, p') - \xi(q, p)\} - 1] \frac{1}{\sqrt{2\pi m}} e^{-\frac{(\rho')^2 + \rho^2}{2m}} \rho(q, p) \rho(q, p') \rho(q, p') \, dp' \, dp. 
\]
Since this expression is symmetric in \( p \) and \( p' \), we can write it as
\[
\Psi^*(\rho, \xi) = \gamma \int \int \left[ \cosh (\xi(q, p') - \xi(q, p)) - 1 \right] \frac{1}{\sqrt{2\pi m}} e^{-\frac{(\rho')^2 + \rho^2}{2m}} \rho(q, p) \rho(q, p') \rho(q, p') \, dp' \, dp. 
\]
For the Vlasov-Fokker-Planck case, we have \( H_2(\rho, \xi) = H(\rho, \xi) - \langle \xi, W(\rho) \rangle \), so that
\[
\mathcal{H}(\rho, \xi) = \langle \xi, G^*(\rho) \rangle + \gamma \int |\nabla_p \xi|^2 d\rho = \langle \xi, F(\rho) \rangle + \gamma \| \xi \|^2_{H_{\rho}}. 
\]
Using that \( dS(\rho) = -\log \rho - 1 - dE(\rho) \) and Lemma 5.2, we calculate \( \Psi^*(\rho, \xi) \):
\[
\Psi^*(\rho, \xi) = H_2 \left( \rho, \xi + \frac{1}{2} dS(\rho) \right) - H_2 \left( \rho, \frac{1}{2} dS(\rho) \right)
= \langle \xi, G^*(\rho) \rangle + \gamma \left( \| \xi + \frac{1}{2} dS(\rho) \|^2_{H_{\rho}} - \| \frac{1}{2} dS(\rho) \|^2_{H_{\rho}} \right)
= \langle \xi, G^*(\rho) \rangle + \int (\nabla_p \xi \cdot \nabla_p dS(\rho)) \, d\rho + \gamma \| \xi \|^2_{H_{\rho}}. 
\]
Following the calculation in (5.8), we find
\[
\Psi^*(\rho, \xi) = \gamma \| \xi \|^2_{H_{\rho}}, 
\]
so that indeed, we reproduce the result of Theorem 5.1.
5.5 Sufficient conditions for a physical extension to GENERIC

In the sections above, the two important ingredients that are shared among the Andersen thermostat and the Vlasov-Fokker-Planck equation are the common structure of $F$ and the reversibility of $G$ with respect to a Gaussian measure. At no point the specific form of $G$ is of importance, except for the calculation of $H_2$. The analysis can therefore be carried out irrespectively of the $G$, if the limit $H_2$ is assumed to exist.

The specific form of $G$ is important, however, from a physical point of view. The assumption is that $G$ models the interaction of the particles with the background solvent. This means that changes only occur under in the momentum variable and not in position. Assuming that the background equilibrates at a faster time scale than that of the evolution of our particles, leads to the assumption of a Gaussian stationary measure.

**Condition 5.4.** The vector-field $\rho \mapsto F(\rho)$ is of the type (5.1) and $G$ and operators $G_n$ defined in terms of $G$ in (5.11), have the following properties:

(a) $G$ acts on the momentum variable only and is independent of the value of $q$,

(b) the operator $G$ interpreted as acting on $p$ only is reversible with respect to a centered Gaussian measure

\[
\rho_m(dp) = \frac{1}{\sqrt{2\pi m}} e^{-\frac{p^2}{2m}} dp,
\]

with $m > 0$.

(c) the limit of

\[
H_2 F(\rho) = \lim_{n \to \infty} \frac{1}{n} e^{-nF(\rho)} (G_n e^{nF}(\rho)
\]

exists for cylinder functions and is of the form $H_2 F(\rho) = \mathcal{H}_2(\rho, dF(\rho))$.

Next, we show that the dynamics in (5.1) can be extended in two ways to GENERIC. For this we use Theorems 4.1 and 4.6. The first theorem can be applied always, i.e. it is the trivial extension to GENERIC, whereas the second result uses that

\[
L(\rho)(dS(\rho) - dE(\rho)) = 0,
\]

see Lemma 5.2.

**Corollary 5.5.** Let Condition 5.4 be satisfied. Then the dynamics

\[
\hat{F}(\rho, e) = \left[ -\text{div}_q (\rho \frac{p}{m}) + \text{div}_p (\rho \nabla_q V + \rho(\nabla_q \Phi * \rho)) + G^*(\rho) \right].
\]

(5.15)

are GENERIC for $\hat{S}(\rho, e) = S(\rho)$, $\hat{E}(\rho, e) = E(\rho)$, $\hat{\Psi}^*((\rho, e), (\xi, r)) = \Psi^*(\rho, \xi)$ and

\[
\hat{L}(\rho, e) \left[ \begin{array}{c} \xi \\ r \end{array} \right] = \left[ \begin{array}{c} r L(\rho) dE(\rho) \\ -(L(\rho) dE(\rho), \xi) \end{array} \right].
\]

**Corollary 5.6.** Let Condition 5.4 be satisfied. Then

\[
F(\rho, e) = \left[ -\text{div}_q (\rho \frac{p}{m}) + \text{div}_p (\rho \nabla_q V + \rho(\nabla_q \Phi * \rho)) + G^*(\rho) \right].
\]

(5.16)

are GENERIC for $\overline{S}(\rho, e) = S(\rho) + (E(\rho) + e)$, $\overline{E}(\rho, e) = E(\rho) + e$, $\overline{\Psi}^*((\rho, e), (\xi, r)) = \Psi^*(\rho, \xi - reD(\rho))$ and

\[
\overline{L}(\rho, e) \left[ \begin{array}{c} \xi \\ r \end{array} \right] = \left[ \begin{array}{c} L(\rho) \xi \\ 0 \end{array} \right].
\]
In both results, note that the different dissipation structure of the equation only affects the dissipation potential $\Psi^*$ of the GENERIC representation.

Remark 5.7. The second result, Corollary 5.6, applied to the Vlasov-Fokker-Planck equation reproduces the result of [3, Section 3.2] after working out the effect of the quadratic potential $\Psi^*$ onto the variables $\xi$ and $r$.

6 Conclusion

In this paper we uncovered a new connection between (a) symmetry properties of Markov processes, expressed through their large-deviation Hamiltonians, and (b) the variational structure of the limiting deterministic evolutions, expressed in terms of GENERIC properties. This connection extends the existing connection for reversible Markov processes and gradient flows, and gives a microscopic basis to the non-quadratic version of GENERIC that we use and that has been identified before (see e.g. [5, 6]).

In the process we also achieved a number of other advances. We gave a formulation of this non-quadratic version of GENERIC in terms of the Hamiltonians that appear naturally in large-deviation theory (Theorem 3.6 and Corollary 3.7). We illustrated this connection by showing how it leads to GENERIC systems for two paradigmatic examples, the Vlasov-Fokker-Planck equation and the Andersen thermostat (Section 5).

Finally, the large-deviation context identified ‘pre GENERIC’ as a class of systems that arises naturally from large-deviation considerations, in the same way as (generalized) gradient flows arise naturally from large-deviation properties of reversible Markov processes. We showed that pre GENERIC and GENERIC are equivalent, in the sense that a system of one type can be converted into a system of the other type. For the implication GENERIC $\implies$ pre GENERIC this is a question of definition; for pre GENERIC $\implies$ GENERIC it turns out that a trivial modification converts any pre GENERIC system into a GENERIC system (Theorem 4.1). This leads to questions of physical interpretation about the relationship between the two concepts, which we leave to future publications.

A Verification of the Jacobi identity in Theorem 4.1

We consider a state space $\hat{Z} := Z \times \mathbb{R}$ with variables $(z, e) \in Z \times \mathbb{R}$. The tangent and cotangent spaces equal $T_{z,e}\hat{Z} = T_zZ \times \mathbb{R}$ and $T^*_{z,e}\hat{Z} = T_zZ \times \mathbb{R}$.

Proposition A.1. Let $W$ be some vector field. Denote by

$$\tilde{L}(z, e) \left( \begin{array}{c} \xi \\ r \end{array} \right) = \left( \begin{array}{c} rW(z) \\ -(W(z), \xi) \end{array} \right) = \left( \begin{array}{cc} 0 & W(z) \\ -W(z) & 0 \end{array} \right) \left( \begin{array}{c} \xi \\ r \end{array} \right).$$

The map $\tilde{L}$ satisfies the Jacobi identity. In other words, the bracket $\{F, G\} := (dF, \tilde{L}dG)$, (applied to smooth functions $F, G : Z \times \mathbb{R} \to \mathbb{R}$) satisfies:

$$\{\{F_1, F_2\}, F_3\} + \{\{F_3, F_1\}, F_2\} + \{\{F_2, F_3\}, F_1\} = 0.$$

The proof of the proposition uses the following auxiliary lemma.
Lemma A.2. For any three pairs \((a_1, b_1), (a_2, b_2), (a_3, b_3) \in \mathbb{R}^2\), denote
\[
\begin{bmatrix}
a_1 \\
b_1 \\
a_2 \\
b_2 \\
a_3 \\
b_3
\end{bmatrix} := a_2a_3b_1 - a_1a_3b_2.
\]
Then we have
\[
\begin{bmatrix}
a_1 \\
b_1 \\
a_2 \\
b_2 \\
a_3 \\
b_3
\end{bmatrix} + \begin{bmatrix}
(a_3) \\
(b_3) \\
(a_1) \\
(b_1) \\
(a_2) \\
(b_2)
\end{bmatrix} + \begin{bmatrix}
(a_2) \\
(b_2) \\
a_1 \\
b_1
\end{bmatrix} = 0. \tag{A.1}
\]
Proof. By definition, the left-hand side of (A.1) equals
\[
a_2a_3b_1 - a_1a_3b_2 + a_1a_2b_3 - a_3a_2b_1 + a_3a_1b_2 - a_2a_1b_3
\]
which indeed equals 0.

Proof of Proposition A.7. The differentials \(dF_i\) consist of two components, one on \(Z\), and one on \(\mathbb{R}\). We denote these components by \(d_zF_i\) and \(d_eF_i\). We introduce the following two sets of functions \(G_{ij}, H_i: Z \times \mathbb{R} \to \mathbb{R}\) defined by
\[
G_{ij}(z, e) := \langle W(z), d_zF_i(z, e) \rangle \cdot d_eF_j(z, e),
H_i(z, e) := \langle W(z), d_zF_i(z, e) \rangle.
\]
First note that \(\{F_i, F_j\} = G_{ij} - G_{ji}\). Next, we calculate \(d_z\) and \(d_e\) of \(G_{ij}\):
\[
d_zG_{ij} = d_zH_i \cdot d_eF_j + H_i \cdot d_zd_eF_j,
\]
\[
d_eG_{ij} = d_eH_i \cdot d_eF_j + H_i \cdot d_e^2d_eF_j
\]
\[
= \langle W, d_zd_eF_i \rangle \cdot d_eF_j + H_i \cdot d_e^2d_eF_j.
\]
Using these differentials, a straightforward calculation yields
\[
\{\{F_i, F_j\}, F_k\} = \begin{bmatrix}
d_z(G_{ij} - G_{ji}) \\
d_e(G_{ij} - G_{ji}) \\
(d_eF_k \cdot W) \\
-H_k
\end{bmatrix}
\]
\[
= d_eF_j \cdot d_eF_k \cdot \langle W, d_zH_i \rangle - d_eF_i \cdot d_eF_k \cdot \langle W, d_zH_j \rangle + H_i \cdot d_eF_k \cdot \langle W, d_zd_eF_j \rangle - H_j \cdot d_eF_k \cdot \langle W, d_zd_eF_i \rangle
\]
\[
- d_eF_j \cdot H_k \cdot \langle W, d_zd_eF_i \rangle + d_eF_i \cdot H_k \cdot \langle W, d_zd_eF_j \rangle
\]
\[
= \begin{bmatrix}
d_eF_i \\
\langle W, d_zH_i \rangle \\
\langle W, d_zd_eF_i \rangle \\
H_i
\end{bmatrix} \cdot \begin{bmatrix}
d_eF_j \\
\langle W, d_zH_j \rangle \\
\langle W, d_zd_eF_j \rangle \\
H_j
\end{bmatrix} + \begin{bmatrix}
H_i \\
d_eF_i \\
\langle W, d_zd_eF_i \rangle \\
H_k
\end{bmatrix} \cdot \begin{bmatrix}
d_eF_j \\
\langle W, d_zH_j \rangle \\
\langle W, d_zd_eF_j \rangle \\
H_j
\end{bmatrix}
\]
\[
+ (H_i \cdot d_eF_k + H_k \cdot d_eF_i) \cdot \langle W, d_zd_eF_j \rangle - (H_j \cdot d_eF_k + H_k \cdot d_eF_j) \cdot \langle W, d_zd_eF_i \rangle.
\]
It follows by Lemma A.2 that
\[
\{\{F_1, F_2\}, F_3\} + \{\{F_3, F_1\}, F_2\} + \{\{F_2, F_3\}, F_1\} = 0.
\]

References

[1] Hans C. Andersen. “Molecular dynamics simulations at constant pressure and/or temperature”. In: The Journal of Chemical Physics 72.4 (1980), pp. 2384–2393. eprint: https://doi.org/10.1063/1.439486.

[2] L. Bertini, D. Gabrielli, and J. L. Lebowitz. “Large Deviations for a Stochastic Model of Heat Flow”. In: Journal of Statistical Physics 121.5 (2005), pp. 843–885.

[3] Manh Hong Duong, Mark A Peletier, and Johannes Zimmer. “GENERIC formalism of a Vlasov-Fokker-Planck equation and connection to large-deviation principles”. In: Nonlinearity 26.11 (2013), pp. 2951–2971.

[4] Jin Feng and Thomas G. Kurtz. Large Deviations for Stochastic Processes. American Mathematical Society, 2006.

[5] Miroslav Grmela. “Fluctuations in extended mass-action-law dynamics”. In: Physica D: Nonlinear Phenomena 241.10 (2012), pp. 976–986.

[6] Miroslav Grmela. “Geometry of multiscale nonequilibrium thermodynamics”. In: Entropy 17.9 (2015), pp. 5938–5964.

[7] Miroslav Grmela and Hans Christian Öttinger. “Dynamics and thermodynamics of complex fluids. I. Development of a general formalism”. In: Phys. Rev. E 56 (6 Dec. 1997), pp. 6620–6632.

[8] Miroslav Grmela and Hans Christian Öttinger. “Dynamics and thermodynamics of complex fluids. I. Development of a general formalism”. In: Physical Review E 56.6 (1997), p. 6620.

[9] C. Kipnis, S. Olla, and S. R. S. Varadhan. “Hydrodynamics and large deviation for simple exclusion processes”. In: Communications on Pure and Applied Mathematics 42.2 (1989), pp. 115–137.

[10] Claude Kipnis and Stefano Olla. “Large deviations from the hydrodynamical limit for a system of independent Brownian particles”. In: Stochastics and Stochastic Reports 33.1-2 (1990), pp. 17–25.
[11] Richard C. Kraaij, Alexandre Lazarescu, Christian Maes, and Mark Peletier. “Deriving GENERIC from a generalized fluctuation symmetry”. In: *Journal of Statistical Physics* (2017). arXiv: [1706.10115](https://arxiv.org/abs/1706.10115).

[12] A. Mielke. “Formulation of thermoelastic dissipative material behavior using GENERIC”. In: *Continuum Mechanics and Thermodynamics* 23.3 (2011), pp. 233–256.

[13] 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”. In: *Potential Analysis* 41.4 (2014), pp. 1293–1327.

[14] L. Onsager. “Reciprocal relations in irreversible processes I & II.” In: *Physical Review* 37:405–426 and 38:2265–2279 (1931).

[15] L. Onsager and S. Machlup. “Fluctuations and Irreversible Processes”. In: *Physical Review* 91.6 (1953), pp. 1505–1512.

[16] H. C. Öttinger. *Beyond Equilibrium Thermodynamics*. Wiley-Interscience, 2005.

[17] H. C. Öttinger. *Beyond Equilibrium Thermodynamics*. Wiley-Interscience, 2005.

[18] Hans Christian Öttinger and Miroslav Grmela. “Dynamics and thermodynamics of complex fluids. II. Illustrations of a general formalism”. In: *Physical Review E* 56.6 (1997), p. 6633.

[19] Hans Christian Öttinger and Miroslav Grmela. “Dynamics and thermodynamics of complex fluids. II. Illustrations of a general formalism”. In: *Physical Review E* 56.6 (1997), p. 6633.

[20] M. A. Peletier, F. Redig, and K. Vafayi. “Large deviations in stochastic heat-conduction processes provide a gradient-flow structure for heat conduction”. In: *Journal of Mathematical Physics* 55.9 (2014), p. 093301.