

Beyond Lagrangians: Noether's theorem in gradient flow PDEs

Nicholas C. White *,†



(Received 16 August 2025; accepted 25 November 2025; published 22 December 2025)

Noether's theorem is traditionally applied to Lagrangian systems to identify conserved quantities. In this work, we apply Noether's theorem instead to a broad class of non-Lagrangian gradient flow partial differential equations (PDEs) arising in physics, showing how continuous symmetries constrain the evolution of such systems and, in certain special cases, still give rise to conserved quantities. We demonstrate symmetry-induced evolutionary constraints numerically on the thin-film equation with capillary and van der Waals forces, and also theoretically derive a conserved quantity for a singular fast-diffusion equation. These results not only provide a tool to analyze gradient flow PDEs; they also demonstrate the utility of Noether's theorem beyond Lagrangians.

DOI: [10.1103/PhysRevE.112.064214](https://doi.org/10.1103/PhysRevE.112.064214)

I. INTRODUCTION

It is a well-known result of multivariable calculus that a scalar quantity's gradient vector is orthogonal to its level sets (see, e.g., Ref. [1]). This fact is the basis for discrete gradient descent techniques such as the Newton-Raphson method, which sequentially take steps in the direction of the gradient to seek local minima. In the limit of infinitesimally short steps, Newton-Raphson descent can be described as an ordinary differential equation (ODE) in gradient flow form [2], in which case the evolution progresses in a direction orthogonal to the level sets.

Generalizing such gradient flow ODEs from finite dimensions to infinite-dimensional functional spaces allows the definition of gradient flow partial differential equations (PDEs), which evolve by following gradients of a governing functional $\mathcal{F}[u]$. The simplest class of such PDEs can be written as $\partial_t u = \delta_u \mathcal{F}[u]$, where δ_u represents the functional derivative; the diffusion equation, for example, falls into this class. A richer and more general class of gradient flow PDEs can be expressed in the weak form $\langle \partial_t u, w \rangle_\beta = \langle \delta_u \mathcal{F}[u], w \rangle_{L^2}$, where $\langle \cdot, \cdot \rangle_\beta$ is some inner product which need not be the L^2 inner product [3]. Examples in physics include the porous medium equation [3], the equation for capillary flow in an interior corner and its variants [4], the foam drainage equation [5], and the thin-film equation and its variants [6,7], all of which are expressible in gradient flow form with the H^{-1} or similar inner products.

It is natural, then, to ask whether the intuition of evolution orthogonal to level sets can be extended to the case of PDE gradient flow and, if so, under what conditions and what

definition of orthogonality. It turns out to be possible by identifying level sets as continuous symmetries, and the evolution as orthogonal to the symmetry characteristic, a result which can be shown using Noether's theorem. Indeed, this connection between gradient flow and Noether's theorem has been recently demonstrated in the context of gradient flow ODEs in neural networks [8,9]. In the present work, we similarly apply Noether's theorem to show how continuous symmetries constrain the evolution for a more general class of gradient flow PDEs, and we further demonstrate applications to physics.

The application of Noether's theorem to gradient flow equations differs from its typical use for conservation laws in Lagrangian and related systems. The one-to-one correspondence between the generalized symmetries of Lagrangian systems and conservation laws was first discovered by Noether [10], in a foundational result of modern mathematical physics, and, since that time, analogous results have been produced by studying the role played in other partial differential equations by continuous symmetries. Olver [11] found the relationship between generalized symmetries and conserved quantities of Hamiltonian systems, a result which was subsequently extended to a larger class of skew-symmetric evolution equations by Strampp [12]. Fokas [13] demonstrated that single-variable evolution equations of the form $\partial_t u(x, t) = K[u(x, t)]$, where $dK[u]/dx$ is the Euler-Lagrange equation for some Lagrangian, also have an exact correspondence between their generalized symmetries and conserved quantities. In contrast with the above results, gradient flow PDEs are not Lagrangian or Hamiltonian (or in the related classes of Strampp and Fokas), and the relevant continuous symmetries are not of the full PDE, but only of the functional driving the gradient flow. Furthermore, the focus is not on the correspondence between symmetries and conserved quantities, but between symmetries and functions orthogonal to the evolution of the system, although in certain special cases continuous symmetries can induce conserved quantities in gradient flow PDEs as well.

One practical application of Noether's theorem in Lagrangian systems is to verify numerical simulations. After performing a simulation, one can confirm whether it comes

*Contact author: ncw@alumni.princeton.edu

†Unaffiliated.

sufficiently close to satisfying conservation of, e.g., energy or angular momentum, and then refine the mesh or change the simulation method as necessary. Similarly, we propose using the fact that gradient flows must evolve orthogonally to their functionals' symmetries as a measure of the effectiveness of numerical schemes. For example, suppose one is computing the evolution of a gradient flow equation whose governing functional satisfies rotational symmetry on a circular domain. If the meshing or discretization of the domain is too coarse, then one may be concerned about the numerical error due to the discretization asymmetry. Computing the numerical solution's orthogonality to the relevant rotational symmetry characteristic provides one metric for that error.

The connection between Noether's theorem and gradient flow ODEs has recently been identified and applied in the field of machine learning, where the set of parameters $\vec{\theta}$ describing a neural network with fixed training data $\vec{x}$ are optimized by gradient descent on a loss function $\ell(\vec{\theta}, \vec{x})$, a process which can be approximated by the continuous gradient flow ODE $\partial_t \vec{\theta} = \nabla_{\vec{\theta}} \ell(\vec{\theta}, \vec{x})$ [8,9]. This connection has been used in particular to understand conserved quantities in neural network training, e.g., the observation of Du *et al.* [14] that the difference between the squared norm of weights (parameters) of subsequent layers of neural networks remains constant during training. Gluch and Urbanke [8] and Kunin *et al.* [9] take the general approach of starting with the symmetries of a system, then applying Noether's theorem to show that the gradient flow evolution is orthogonal to those symmetries and noting the induced conserved quantities where applicable; we take a similar route in this paper. Marcotte *et al.* [15] instead consider conserved quantities of neural networks directly and prove that the number of such quantities is related to the dimension of the Lie algebra generated by a reparametrization of the dependent variable $\vec{\theta}$; they note the parallels to Noether's theorem without invoking it directly.

Gluch and Urbanke [8] and Marcotte *et al.* [16] further generalized their results beyond gradient flow to inertial equations with second-order time derivatives; while equations of the form $\partial_t^2 u = \delta_u \mathcal{F}[u]$ are expressible as Lagrangians and are the subject of the original Noether's theorem, the addition of first-order damping terms in general makes such equations no longer expressible in Lagrangian form, necessitating a different approach. Although the aforementioned works explicitly discussed only unweighted inner products, Gluch and Urbanke wrote a very general form of gradient-flow-like equation, $E(t, \vec{\theta}, \partial_t \vec{\theta}, \partial_t^2 \vec{\theta}) = \nabla_{\vec{\theta}} \ell(\vec{\theta})$ which effectively includes the possibility of a gradient flow ODE with a weighted inner product.

In this work, we study gradient flow equations which are first order in time, and rather than multivariable ODEs, we instead consider scalar gradient flow PDEs of the type described by Otto [3] with a general class of inner products, including in particular a variety of diffusion-like equations arising in physics. While the general concept is similar to the ODE case, the boundary conditions and nontrivial inner products that can arise in PDEs complicate the results. We also determine a condition under which conserved quantities arise in the general case, and in special cases (including ones considered by Kunin *et al.*), further clarify the condition as equivalent

to self-adjointness of the symmetry characteristic's Fréchet derivative. We demonstrate the application of the result to a liquid thin film equation driven by capillary and van der Waals forces, and show how it can be used as a metric of the numerical method. Finally, we demonstrate how the scaling symmetry of a singular fast-diffusion equation produces a conserved quantity.

The paper is organized as follows: In Sec. II, we review definitions of continuous symmetries and Noether's theorem. Section III will relate the main theorem regarding evolutionary constraints, as well as corollaries describing the special conditions for conserved quantities to arise. Section IV demonstrates the results with two examples: first, a numerical computation of the evolutionary constraints of a liquid thin film equation; second, a derivation of the symmetry-induced conserved quantity in a singular fast-diffusion equation.

II. BACKGROUND

We begin by reviewing some definitions, and the statement of Noether's theorem, following Olver [17]. For the duration of this work, we consider functionals of scalar functions $u : \Omega \rightarrow \mathbb{R}$, with domain $\Omega \subseteq \mathbb{R}^k$. Furthermore, we in general assume homogeneous boundary conditions on $\partial\Omega$ in order to avoid boundary terms (boundary terms will be explicitly shown for specific examples in Sec. IV). For brevity, details such as smoothness requirements and generalizations to higher dimensions will be omitted. All functions will be assumed to be sufficiently smooth, and the interested reader is referred to Olver [17] for a more rigorous approach.

A. Variational derivatives and the Euler operator

Given a functional

$$\mathcal{F}[u] = \int_{\Omega} \mathcal{F}[u] d^k x, \quad (1)$$

with square brackets denoting that $\mathcal{F}$ may depend on u , ∇u , $\nabla^2 u$, etc., its *variational derivative* with respect to a function u will be denoted $\delta_u \mathcal{F}$, defined by satisfying

$$\int_{\Omega} \delta_u \mathcal{F}[u] \cdot m d^k x = \left(\frac{d}{d\varepsilon} \mathcal{F}[u + \varepsilon m] \right)_{\varepsilon=0} \quad (2)$$

for sufficiently smooth functions m with compact support satisfying the boundary conditions imposed on u [17]. Note that $\delta_u \mathcal{F}$ is a vector with the same dimension as u ; this work will assume u is a scalar, so $\delta_u \mathcal{F}$ will also be a scalar.

Let J denote the multi-index $J = \{j_1, j_2, \dots, j_k\}$, each j_i being a non-negative integer, and let $|J| = \sum_i j_i$ be the size of J . Let d_J denote the differential operator $d_J = d^{j_1+j_2+\dots+j_k} / dx_1^{j_1} dx_2^{j_2} \dots dx_k^{j_k}$. Then the *Euler operator* is defined as

$$\begin{aligned} E_u &= \sum_J (-1)^{|J|} d_J \frac{\partial}{\partial (d_J u)} \\ &= \frac{\partial}{\partial u} - \sum_{i=1}^k \frac{d}{dx_i} \frac{\partial}{\partial (du/dx_i)} \\ &\quad + \sum_{i=1}^k \sum_{\ell=1}^k \frac{d^2}{dx_i dx_{\ell}} \frac{\partial}{\partial (d^2 u / dx_i dx_{\ell})} - \dots, \end{aligned} \quad (3)$$

where the sum in the first expression extends over all sets of multi-indices J [17]. Note in particular that

$$\delta_u \mathcal{F} = E_u[\mathcal{F}]; \quad (4)$$

that is, the variational derivative of a functional is equal to the Euler operator applied to the functional density [17]. Furthermore, if $\mathcal{F}$ is a Lagrangian (and hence $\mathcal{F}$ is a Lagrangian density), then $E_u[\mathcal{F}]$ is the corresponding Euler-Lagrange equation.

B. Symmetries and characteristics

A generalized variational symmetry group of a functional $\mathcal{F} = \int_{\Omega} \mathcal{F}[u] d^k x$ with $\bar{x} \in \Omega$ and $u(x) \in U$ is a group of local transformations of x and u that leave the functional invariant up to the integral of a total derivative (i.e., invariant up to boundary terms). Specifically, $g : \Omega \times U \rightarrow \Omega \times U$ such that

$$\begin{aligned} \int_{\Omega_1} \mathcal{F}[g \cdot u] d^k g \cdot x &= \int_{\Omega_1} \mathcal{F}[u] d^k x + \int_{\Omega_1} \nabla \cdot \vec{A}[u] d^k x \\ &= \int_{\Omega_1} \mathcal{F}[u] d^k x + \int_{\partial\Omega_1} \vec{A}[u] \cdot \hat{n} d^{k-1} x \end{aligned} \quad (5)$$

for all $\Omega_1 \subset \Omega$ (note that defining the action of g on u and x is sufficient to determine its action on derivatives of u , via the prolongation equation) [17].

Such a symmetry group is a Lie group and hence may be described in terms of a set of infinitesimal generators

$$\vec{v} = \sum_{i=1}^k \xi^i[x, u] \frac{\partial}{\partial x_i} + \phi[x, u] \frac{\partial}{\partial u} \quad (6)$$

satisfying

$$\vec{v}(\mathcal{F}) + \mathcal{F} \nabla \cdot \xi + \nabla \cdot \vec{A} = 0, \quad (7)$$

where ξ^i , ϕ , and $\vec{A}$ may depend both on u and on derivatives of u [17].

While the inclusion of the total derivative term implies that a generalized variational symmetry does not necessarily leave a functional $\mathcal{F}$ invariant, it does leave the Euler-Lagrange equation of that functional invariant. That is, it adds a term to $\mathcal{F}$, which is in the kernel of E_u [17]. For the remainder of this work, the term *variational symmetry* will be used to refer to generalized variational symmetries.

The characteristic of a symmetry generated by $\vec{v}$ as in Eq. (6) is given by

$$Q[u] = \phi[x, u] - \sum_{i=1}^k \xi^i[x, u] \frac{\partial u}{\partial x_i}. \quad (8)$$

Note that a specific solution u_{spec} is invariant under the group generated by $\vec{v}$ if and only if $Q[u_{\text{spec}}] = 0$ [18]. Except where necessary, functionals such as $Q[u]$ and $\mathcal{F}[u]$ will be written as Q and $\mathcal{F}$, without brackets, with their potential dependence on u , ∇u , etc. implied.

C. Noether's theorem

Theorem 1 (Noether's theorem). A functional $\mathcal{L}[u] = \int \mathcal{L} d^k x$ has a variational symmetry with characteristic Q if and

only if $\exists \vec{B} : \Omega \rightarrow \mathbb{R}^k$ such that $E_u[\mathcal{L}]Q = \nabla \cdot \vec{B}$. In particular, when $E_u[\mathcal{L}]$ vanishes, then $\nabla \cdot \vec{B} = 0$, i.e., $\vec{B}$ is a conserved flux [17].

The typical application of Noether's theorem in physics is to Lagrangian problems, in which $\mathcal{L} = \int \mathcal{L} d^n x dt$ is a Lagrangian integrated over both space and time. On shell, $E_u[\mathcal{L}] = 0$, so that $0 = \nabla \cdot \vec{B} = \partial_t B^{(t)} + \nabla_i B^{(i)}$, where we have separated the time and space components of $\vec{B}$. Integrating the result over a time-independent spatial domain yields $\partial_t \int B^{(t)} d^n x = 0$. This last result holds only if the boundary conditions are such that the boundary terms vanish.¹ In the derivation below, we similarly ignore boundary conditions, assuming they are vanishing; boundary conditions will be discussed in the specific examples in Sec. IV.

III. NOETHER'S THEOREM APPLIED TO GRADIENT FLOW

In the following results, $\langle \cdot, \cdot \rangle$ (no subscript) will denote the L^2 inner product, i.e., $\langle u, v \rangle = \int uv d^k x$, while $\langle \cdot, \cdot \rangle_{\mathcal{G}}$ will denote an unspecified inner product. For example, $\langle \cdot, \cdot \rangle_{\mathcal{G}}$ may be the L^2 inner product, a weighted L^2 inner product $\langle \cdot, w \cdot \rangle$, the H^{-1} inner product $\langle \cdot, \cdot \rangle_{H^{-1}} = \langle -\nabla^{-2} \cdot, \cdot \rangle$, where ∇^{-2} is the inverse Laplacian operator, etc.

We also define the inner product $\langle \cdot, \cdot \rangle_{\mathcal{G}^{-1}}$ as

$$\langle u, v \rangle_{\mathcal{G}^{-1}} = \langle \mathcal{G}u, v \rangle = \int_{\Omega} s v d^k x, \quad \text{with } \mathcal{G}s = u, \quad (9)$$

where $\mathcal{G}$ is a self-adjoint and surjective linear operator with a rule for choosing an inverse. In particular, we will be interested in cases where $\mathcal{G}$ is an elliptic differential operator. When $\mathcal{G}$ is a differential operator, it must be paired with appropriate smoothness constraints and boundary conditions to enable a unique choice of inverse. For example, $\mathcal{G}$ may be $-\nabla^2$, denoting the negative of the scalar Laplacian operator, in which case $\langle \cdot, \cdot \rangle_{\mathcal{G}^{-1}} = \langle \cdot, \cdot \rangle_{H^{-1}}$; or it may be a different elliptic operator [3].

A gradient flow PDE will be defined following [3], as an equation of the form

$$\langle \partial_t u, w \rangle_{\beta} = \langle \delta_u \mathcal{F}, w \rangle. \quad (10)$$

A. Evolutionary constraints of gradient flow PDEs

Proposition 1. Let $\vec{v}[u]$ be a variational symmetry of the functional $\mathcal{F}[u] = \int_{\Omega} \mathcal{F}[u] d^k x$, with symmetry characteristic $Q[u]$. If u is governed by a gradient flow equation $\langle \partial_t u, w \rangle_{\beta} = \langle \delta_u \mathcal{F}, w \rangle$, then $\partial_t u$ is orthogonal to Q under the $\langle \cdot, \cdot \rangle_{\beta}$ inner product, $\langle \partial_t u, Q \rangle_{\beta} = 0$, given appropriate boundary conditions.

Proof. Let $w = Q$. By Noether's theorem, $\langle \delta_u \mathcal{F}, Q \rangle = \int_{\Omega} E_u[\mathcal{F}] Q d^k x = \int_{\Omega} \nabla \cdot \vec{B} d^k x = \oint_{\partial\Omega} \vec{B} \cdot d\hat{n}$, where $\hat{n}$ is the normal vector. If the boundary conditions are such that $\oint_{\partial\Omega} \vec{B} \cdot d\hat{n}$ vanishes, then $\langle \delta_u \mathcal{F}, Q \rangle = 0$, and hence $\langle \partial_t u, Q \rangle_{\beta} = 0$. ■

¹There are cases where the boundary term cannot vanish, such as when $\vec{B}$ is independent of u . In that case, one obtains a quantity that varies according to some u -independent function of t , rather than a quantity that is constant in time.

One interesting aspect of this result is that it reverses the usual importance of the terms in Noether's theorem. In Lagrangian systems, where $E_u[\mathcal{L}]Q = \nabla \cdot \vec{B}$, the $E_u[\mathcal{L}]Q$ term vanishes on-shell and it is the $\vec{B}$ term that is nontrivial and describes the conserved quantity. But for gradient flow, it is the $\vec{B}$ term that vanishes and the $E_u[\mathcal{F}]Q$ term that is relevant to the final result.

In the finite-dimensional case, we know that the gradient of a scalar function f is orthogonal to contours of constant f . In particular, a symmetry of f defines contours of constant f , and so the gradient must be orthogonal to the symmetry. In the infinite-dimensional gradient flow, Proposition 1 defines the sense in which the evolution is orthogonal to any variational symmetry.

Note that the ODE version of Proposition 1 was also shown by Gluch and Urbanke [8] and Kunin *et al.* [9] in the context of neural networks. In Sec. IV, we demonstrate how the generalization to PDEs with a variety of inner products enables the application of this result to a much wider class of equations arising in physics.

B. Conserved quantities

The evolutionary constraints do not appear to always lead to conserved quantities, although we can prove a sufficient (but perhaps not necessary) condition for conserved quantities to arise.

Corollary 1. Let $\vec{v}[u]$ be a variational symmetry of the functional $\mathcal{F}[u]$, with symmetry characteristic $Q[u]$. Let u be governed by a gradient flow equation $\langle \partial_t u, w \rangle_{\mathcal{G}^{-1}} = \langle \delta_u \mathcal{F}, w \rangle$. If $Q = \mathcal{G} \delta_u \mathcal{R}$ for some $\mathcal{R} = \int \mathcal{R} d^k x$, then $\mathcal{R}$ is a conserved quantity under the gradient flow, i.e., $\partial_t \mathcal{R} = \partial_t \int \mathcal{R} d^k x = 0$, given appropriate boundary conditions.

Proof. It was established in Proposition 1 that $\langle \partial_t u, Q \rangle_{\mathcal{G}^{-1}} = 0$. Hence, $0 = \langle \partial_t u, Q \rangle_{\mathcal{G}^{-1}} = \langle \partial_t u, \mathcal{G} \delta_u \mathcal{R} \rangle_{\mathcal{G}^{-1}} = \langle \partial_t u, \delta_u \mathcal{R} \rangle = \partial_t \int \mathcal{R} d^k x$. ■

Intuitively, this condition can be understood as requiring the existence of a different functional landscape that swaps the roles of Q and $\partial_t u$. In the $\mathcal{F}$ landscape, $\partial_t u$ follows the gradient (the steepest direction), and is orthogonal to Q , the contour lines of constant value. Similarly, if there is some $\mathcal{R}$ landscape of which Q is the gradient, then the orthogonal $\partial_t u$ must follow contour lines, i.e., $\mathcal{R}$ is constant along every path of evolution of u .

The question of when $Q = \mathcal{G} \delta_u \mathcal{R}$ then arises. We do not expect $\mathcal{G}^{-1}Q$ to always be a functional derivative, just as, in the finite-dimensional case, not every vector field is a gradient of some scalar. Before proceeding, note that the Fréchet derivative of a function (or functional) $\mathcal{F}[u]$ is the differential operator defined such that its action on a function m is given by

$$\mathbf{D}_{\mathcal{F}}(m) = \left(\frac{d}{d\varepsilon} \mathcal{F}[u + \varepsilon m] \right)_{\varepsilon=0}. \quad (11)$$

Corollary 2. Let $\vec{v}[u]$ be a variational symmetry of the functional $\mathcal{F}[u]$, with characteristic $Q[u]$, with $u(\vec{x})$ in a totally star-shaped subdomain of $X \times U$. Let u be governed by a gradient flow equation $\langle \partial_t u, w \rangle_{\mathcal{G}^{-1}} = \langle \delta_u \mathcal{F}, w \rangle$. Further

assume that $\mathcal{G}^{-1}Q$ is expressible without integral operators.² If the Fréchet derivative $\mathbf{D}_{\mathcal{G}^{-1}Q}$ is self-adjoint under the L^2 inner product, then $Q = \mathcal{G} \delta_u \mathcal{R}$ for some $\mathcal{R} = \int \mathcal{R} d^k x$, and $\mathcal{R}$ is a conserved quantity under the gradient flow, given appropriate boundary conditions.

In the special case that $\mathcal{G}$ is independent of u , this condition is equivalent to $\mathbf{D}_Q$ being self-adjoint under the $\mathcal{G}^{-1}$ inner product (i.e., $\mathbf{D}_Q \mathcal{G} = \mathcal{G} \mathbf{D}_Q^*$, where $\mathbf{D}_Q$ is the Fréchet derivative of Q and $\mathbf{D}_Q^*$ is its L^2 adjoint).

Proof. It can be shown that an expression M is the variational derivative of some quantity N (i.e., $M = \delta_u N$ for some N) if and only if $\mathbf{D}_M = \mathbf{D}_M^*$, where $\mathbf{D}_M$ is the Fréchet derivative of M and $\mathbf{D}_M^*$ its adjoint under the L^2 inner product³ [17]. Therefore, $Q = \mathcal{G} \delta_u \mathcal{R}$ for some $\mathcal{R}$ if and only if $\mathbf{D}_{\mathcal{G}^{-1}Q} = [\mathbf{D}_{\mathcal{G}^{-1}Q}]^*$.

If $\mathcal{G}$ is independent of u , then it commutes with the Fréchet derivative, and hence the condition becomes $\mathcal{G}^{-1} \mathbf{D}_Q = \mathbf{D}_Q^* \mathcal{G}^{-1}$, or $\mathbf{D}_Q \mathcal{G} = \mathcal{G} \mathbf{D}_Q^*$, where we used the fact that $\mathcal{G}$ is self-adjoint. Equivalently, $Q = \mathcal{G} \delta_u \mathcal{R}$ if and only if $\mathbf{D}_Q$ is self-adjoint under the $\mathcal{G}^{-1}$ inner product, because $\langle u, \mathbf{D}_Q v \rangle_{\mathcal{G}^{-1}} = \langle \mathcal{G}^{-1} u, \mathbf{D}_Q v \rangle_{L^2} = \langle \mathcal{G} \mathbf{D}_Q^* \mathcal{G}^{-1} u, v \rangle_{\mathcal{G}^{-1}}$. ■

This result implies that symmetries of the independent variables x_i (whether scaling symmetries, translational symmetries, rotational symmetries, or others) typically cannot induce conserved quantities in this way, since a symmetry characteristic $\xi^i[x, u] \partial u / \partial x_i$ will not be self-adjoint, except for very specific forms of $\xi^i[x, u]$ containing derivatives of u . Instead, conserved quantities typically arise from symmetries of the dependent variable u .

Note that Corollary 1, if applied to nonscalar ODEs instead of scalar PDEs, would describe the conserved quantities noted by Kunin *et al.* [9] but differs from their stated conservation condition (without contradicting it). Those authors noted that for a symmetry $\vec{Q}$ of the function $\vec{F}(\vec{u})$ governing a gradient flow ODE $\partial_t \vec{u} = \vec{F}(\vec{u})$, if it happens to be the case that $\vec{F} \cdot (\partial_u \vec{Q}) \cdot \vec{u} = 0$, then $\vec{u} \cdot \vec{Q}$ is a quantity conserved in time, where $\cdot$ is the vector dot product. In other words, while their conserved quantity is schematically “ u times Q ,” ours is “the functional integral of Q with respect to u .” These two definitions are equivalent when $\vec{Q} = \text{const}$ or $\vec{Q} = \sum_j A_{ij} u_j$ for matrix A symmetric (corresponding to the self-adjointness condition of Corollary 2), as is the case in all of the examples of Kunin *et al.*, as well as for $\vec{Q}$ proportional to other powers of $\vec{u}$. It may be worth investigating cases of ODEs for which only one or the other definition produces a conserved quantity. For PDEs or nontrivial inner products, the potential dependence of Q and $\mathcal{G}$ on spatial derivatives of u makes identification

²This condition is stronger than necessary, but the exact restrictions on the form of $\mathcal{G}$ are beyond the scope of this paper. Although we have not proven the statement for operators such as $\mathcal{G} = -\nabla^2$, these work in practice, as will be demonstrated in the next section.

³The self-adjointness condition is analogous to finite-dimensional gradient vectors having a self-adjoint Jacobian, and implies a closed differential form; the star-shaped domain requirement is necessary for closed forms to be exact (i.e., variational derivatives of a functional).

of a simple analog of Kunin *et al.*'s condition difficult, and Corollary 1 instead arises more naturally.

IV. EXAMPLES

A. Example 1: Thin film equation with van der Waals forces

1. Gradient flow form and symmetries

Consider the nondimensionalized thin film equation with constant surface tension and van der Waals forces,

$$\partial_t h = -\nabla \cdot \left[\frac{h^3}{3} \nabla \nabla^2 h \right] - A \nabla \left[\frac{1}{h} \nabla h \right] = \nabla \cdot \left[\frac{h^3}{3} \nabla p \right], \quad (12)$$

$$p = -\nabla^2 h + A h^{-3}, \quad (13)$$

where $h(x, y)$ is nondimensional film thickness, A is the nondimensional Hamaker constant, $p(x, y)$ is the effective nondimensional pressure, and ∇^2 is the Laplacian operator. This equation describes a thin liquid film on a flat substrate whose dynamics are driven by surface tension and van der Waals interaction of the liquid-gas interface with the solid substrate (attraction if $A > 0$ and repulsion if $A < 0$) [19].

Let us take a no-flux boundary condition (e.g., due to solid walls) and a contact angle of 90° on the boundary of the domain Ω . Mathematically, this is expressed by the Neumann conditions $\partial_n p|_{\partial\Omega} = 0$ and $\partial_n h|_{\partial\Omega} = 0$, where ∂_n denotes $\hat{n} \cdot \nabla$, the gradient normal to the domain boundary. We further assume that the film thickness h is everywhere positive.

Equation (12) may be expressed in gradient flow form as

$$\langle \partial_t h, w \rangle_{\mathcal{G}^{-1}} = \langle \nabla^2 h - A h^{-3}, w \rangle = \langle \delta_h \mathcal{F}, w \rangle, \quad (14)$$

where

$$\langle v, w \rangle_{\mathcal{G}^{-1}} = \langle s, w \rangle, \quad (15)$$

$$\mathcal{G}s = -\nabla \cdot \left(\frac{h^3}{3} \nabla s \right) = v, \quad \partial_n s|_{\partial\Omega} = 0, \quad (16)$$

(i.e., $\mathcal{G} = -\nabla \cdot [(h^3/3)\nabla]$, a Laplacian-like elliptic operator, and $\langle \rangle_{\mathcal{G}^{-1}}$ is defined based on the solution $\mathcal{G}s = v$ with Neumann boundary conditions), and

$$\mathcal{F} = -\frac{1}{2} \nabla h \cdot \nabla h + \frac{A}{2} h^{-2}. \quad (17)$$

We chose this equation as an example because of its complexity and nonlinearity in both the inner product and the governing functional, emphasizing the broad applicability of the gradient flow symmetry analysis methodology.

Hence, by Proposition 1, if Q is a variational symmetry characteristic of $\mathcal{F}$ then

$$\langle \partial_t h, Q \rangle_{\mathcal{G}^{-1}} = \int_{\Omega} (\partial_t h) s dx dy = 0, \quad (18)$$

$$\text{with } s \text{ defined by } \nabla \cdot \left(\frac{h^3}{3} \nabla s \right) = Q, \quad \partial_n s|_{\partial\Omega} = 0. \quad (19)$$

Just as Noether's theorem in Lagrangian systems provides quantities that are conserved in time only with the appropriate boundary conditions, the symmetry characteristics of the gradient flow functional will be orthogonal to the evolution vector only with the appropriate boundary conditions. In this

case specifically, the required boundary condition on s for the boundary terms to vanish is

$$\int_{\partial\Omega} [(\partial_t h) h^3 \partial_n s - \partial_n (h^3 \partial_t h) s] = 0. \quad (20)$$

Because of our initial choice of Neumann boundary conditions on h , we had to choose Neumann boundary conditions on s as well ($\partial_n s|_{\partial\Omega} = 0$). If we had instead chosen Dirichlet boundary conditions on h , then the homogeneous Dirichlet boundary condition $s|_{\partial\Omega} = 0$ would instead be used.

Performing a variational symmetry analysis on $\mathcal{F}$, we find three symmetries:

(1) $\bar{v}_1 = \partial_x$, corresponding to translation in x . The characteristic is $Q_1 = -\partial_x h$.

(2) $\bar{v}_2 = \partial_y$, corresponding to translation in y . The characteristic is $Q_2 = -\partial_y h$.

(3) $\bar{v}_3 = y\partial_x - x\partial_y$, corresponding to rotation in the x - y plane. The characteristic is $Q_3 = x\partial_y h - y\partial_x h$.

2. Numerical example

We now demonstrate the results with a numerical example, computed using the finite element method (FEM) with FreeFEM [20]. All code and data are available from Ref. [21]. Let us consider, as an example, the rotational symmetry $\bar{v} = \bar{v}_3 = y\partial_x - x\partial_y$. This symmetry is valid if the domain Ω is circular and the boundary conditions are rotationally symmetric (our choice of Neumann boundary conditions $\partial_n h = 0$, $\partial_n p = 0$, is valid). Therefore, the function $Q = (x\partial_y h - y\partial_x h)$ is always orthogonal to the evolution $\partial_t h$ under the $\mathcal{G}^{-1}$ inner product,

$$\langle \partial_t h, x\partial_y h - y\partial_x h \rangle_{\mathcal{G}^{-1}} = 0. \quad (21)$$

Note in particular that the initial condition is *not* required to be rotationally symmetrical for this orthogonality to hold; it is sufficient for the domain shape and boundary condition to satisfy rotational symmetry (in fact, the equation is interesting only when the data are not rotationally symmetric, because $Q = x\partial_y h - y\partial_x h$ vanishes everywhere on rotationally symmetric data).

We numerically solve Eq. (12) with domain Ω being a circle of radius 1, with boundary conditions $\partial_n h|_{\partial\Omega} = \partial_n p|_{\partial\Omega} = 0$. We set the nondimensional Hamaker constant $A = 5$ (making the film interface attracted to the substrate), and choose the initial condition

$$h(t=0) = 1 - 0.2e^{-[\frac{x-0.4}{0.35}]^2 - [\frac{y-0.3}{0.3}]^2} + 0.2e^{-[\frac{x}{0.3}]^2 - [\frac{y}{0.5}]^2},$$

$$\tilde{x} = \frac{x+0.2}{\sqrt{2}} + \frac{y+0.4}{\sqrt{2}}, \quad \tilde{y} = \frac{x+0.2}{\sqrt{2}} - \frac{y+0.4}{\sqrt{2}} \quad (22)$$

[see Fig. 1(d)], with the FEM solver allowed to adjust it at the boundary to conform to the Neumann boundary condition. This initial condition is simply a film of thickness 1, with the addition of a negative Gaussian term (a depression in the film's surface), and a positive Gaussian term (a bump on the film's surface). It is chosen to be rotationally asymmetric so that the symmetry characteristic will evaluate to an interesting nonzero function.

The thin-film simulation with the displayed initial condition was performed using the finite element method three

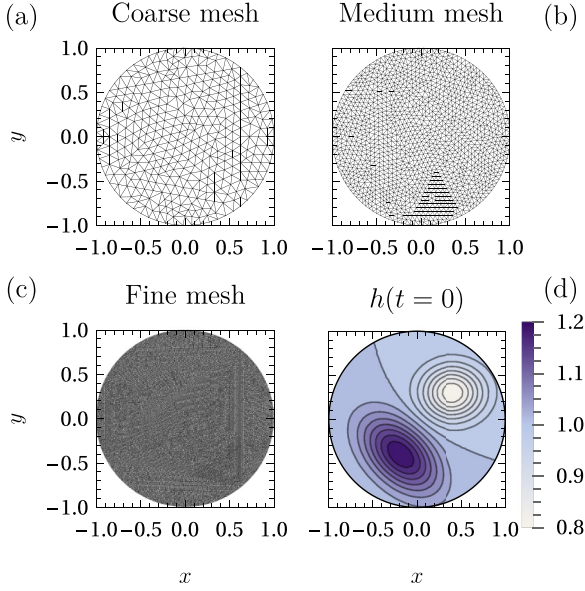


FIG. 1. (a)–(c) Coarse, medium, and fine meshes used in finite element method (FEM) simulations of thin-film equation with constant surface tension and van der Waals forces, Eq. (12), on a circular domain of radius 1. (a) Coarse mesh has 709 elements, with typical side length 0.1. (b) Medium mesh has 2742 elements, with a typical side length 0.05. (c) Fine mesh has 68 763 elements, with a typical side length of 0.01. (d) Initial condition of film thickness h , Eq. (22), composed of two Gaussian functions, one negative (a depression) in the upper right, and one positive (a bump) in the lower left. The initial condition is chosen to be asymmetrical, so that symmetry characteristics will evaluate to nonzero functions.

times: once with a coarse mesh, once with a medium mesh (the results of which were shown in Fig. 2), and once with a fine mesh. The coarse mesh had 709 triangular elements, each with a typical side length of 0.1 [Fig. 1(a)], the medium mesh had 2742 elements, with a typical side length of 0.05 [Fig. 1(b)], and the fine mesh had 68 763 elements, with a typical side length of 0.01 [Fig. 1(c)]. All three meshes were chosen to have elements smaller than the scale of variations in the initial condition. The system was integrated from time $t = 0$ to $t = 0.05$, with step size 0.001, using the backward Euler method. At each step, a fixed-point iteration method was run until the maximum difference between subsequent iterations of h was less than 10^{-9} .

The evolution of Eq. (12) with the example initial condition in a circular domain of radius 1 is shown in Fig. 2; all results in this figure were computed with a mesh with element size 0.05 [labeled “medium mesh” in Fig. 1(b)]. The rows display fluid film thickness h , rotational symmetry characteristic Q , orthogonal function $\mathcal{G}^{-1}Q$, and evolution vector $\partial_t h$. Each column shows a different time in the evolution; the first column is the first time step at $t = 0.001$, the second column displays $t = 0.01$, and the final row displays $t = 0.05$.

Consider first the film thickness h (row 1). It begins with a bump in the lower left and a depression in the upper right; as time progresses, the bump and depression each migrate to the edge of the domain and become wider. The bump decreases in thickness, driven by the flattening effect of the surface tension. The film at the depression initially thickens, also due to the

surface tension, but then thins due to the van der Waals force of the substrate pulling on the film’s surface.

Row 2 of Fig. 2 shows the symmetry characteristic Q . Q would vanish if the data h satisfied the symmetry; that is, if h were rotationally symmetric, then Q would be identically zero. Q thus provides a local measure of the deviation of h from rotational symmetry.

The orthogonal function $\mathcal{G}^{-1}Q$ (row 3) would thus also vanish if h were rotationally symmetric; however, it is not a local descriptor of deviation from the symmetry, but a nonlocal one, because $\mathcal{G}^{-1}$ is a nonlocal operator. Note that, due to the choice of Neumann boundary conditions, $\mathcal{G}^{-1}Q$ can be defined only up to a constant. For the simulation, the choice was arbitrarily made to set the value at $(x, y) = (1, 0)$ to zero. This does not affect the orthogonality with $\partial_t h$; precisely because of the Neumann boundary conditions, volume is conserved and $\int_{\Omega} \partial_t h dx dy = 0$.

As the film evolves, $\partial_t h$ (row 4) is always orthogonal to Q (row 2) under the $\mathcal{G}^{-1}$ inner product; equivalently, $\partial_t h$ is always orthogonal to $\mathcal{G}^{-1}Q$ (row 3) under the L^2 inner product. Again, this must hold even though the initial condition is not rotationally symmetric. As with the application of Noether’s theorem to Lagrangian problems, the symmetry is a quality of the governing equations, not of the data. It can be seen visually that $\partial_t h$ has an approximate (though not exact) discrete symmetry of reflection about the diagonal line $y = x$. To compensate, $\mathcal{G}^{-1}Q$ is approximately odd (up to an overall constant) about that same line.

In ODE gradient flow, it is relatively easy to visualize that the evolution vector must be orthogonal to the level sets of the finite-dimensional landscape governing the gradient flow. Intuition for PDEs is more difficult, but for many gradient flow PDEs arising in physics, such as the thin-film equation, the governing functional can be considered a free energy, and the system evolves to minimize that free energy; the evolution is therefore orthogonal to directions that would leave that free energy constant.

In the present thin-film example, a given fluid interface h_1 , expressed in polar coordinates as $h_1(r, \theta)$, has the same free energy as a rotated version of that same interface, $h_1(r, \theta + \text{const})$. If that film interface were to simply rotate in time, $h(r, \theta, t) = h_1(r, \theta - t + t_1)$, instead of evolving according to Eq. (12), then the motion of that interface, $\partial_t h_1 = -\partial_{\theta} h_1 = Q$ (row 2), would represent evolution along a free-energy level set (the speed of rotation is irrelevant; a different speed would simply introduce a constant multiplicative factor). Thus, the actual thin-film evolution of a film h which has the shape $h(r, \theta, t_1) = h_1(r, \theta)$ at time t_1 must be orthogonal to the evolution of the hypothetical rotating film, i.e., $\partial_t h(r, \theta, t)|_{t=t_1}$ must be orthogonal to $-\partial_{\theta} h(r, \theta, t)|_{t=t_1} = Q|_{t=t_1}$.

For a gradient flow PDE governed by an L^2 inner product, the intuition could end here. But the thin-film equation is governed by the $\mathcal{G}^{-1}$ inner product, where $\mathcal{G}^{-1}$ is the operator that converts between surface motion $\partial_t h$ and pressure p , the latter of which serves as the functional gradient of the free energy. In particular, $\mathcal{G}^{-1} \partial_t h = -p$, and $\mathcal{G}^{-1}(-\partial_{\theta} h) = \mathcal{G}^{-1}Q = -p_{\text{rot}}$ (row 3), a hypothetical pressure (replacing the surface tension and van der Waals pressure) that would drive the film to rotate. The statement $\langle \partial_t h, Q \rangle_{\mathcal{G}^{-1}} = 0$ is thus equivalent to $\langle \partial_t h, p_{\text{rot}} \rangle = 0$ and $\langle \partial_{\theta} h, p \rangle = 0$: the actual film surface

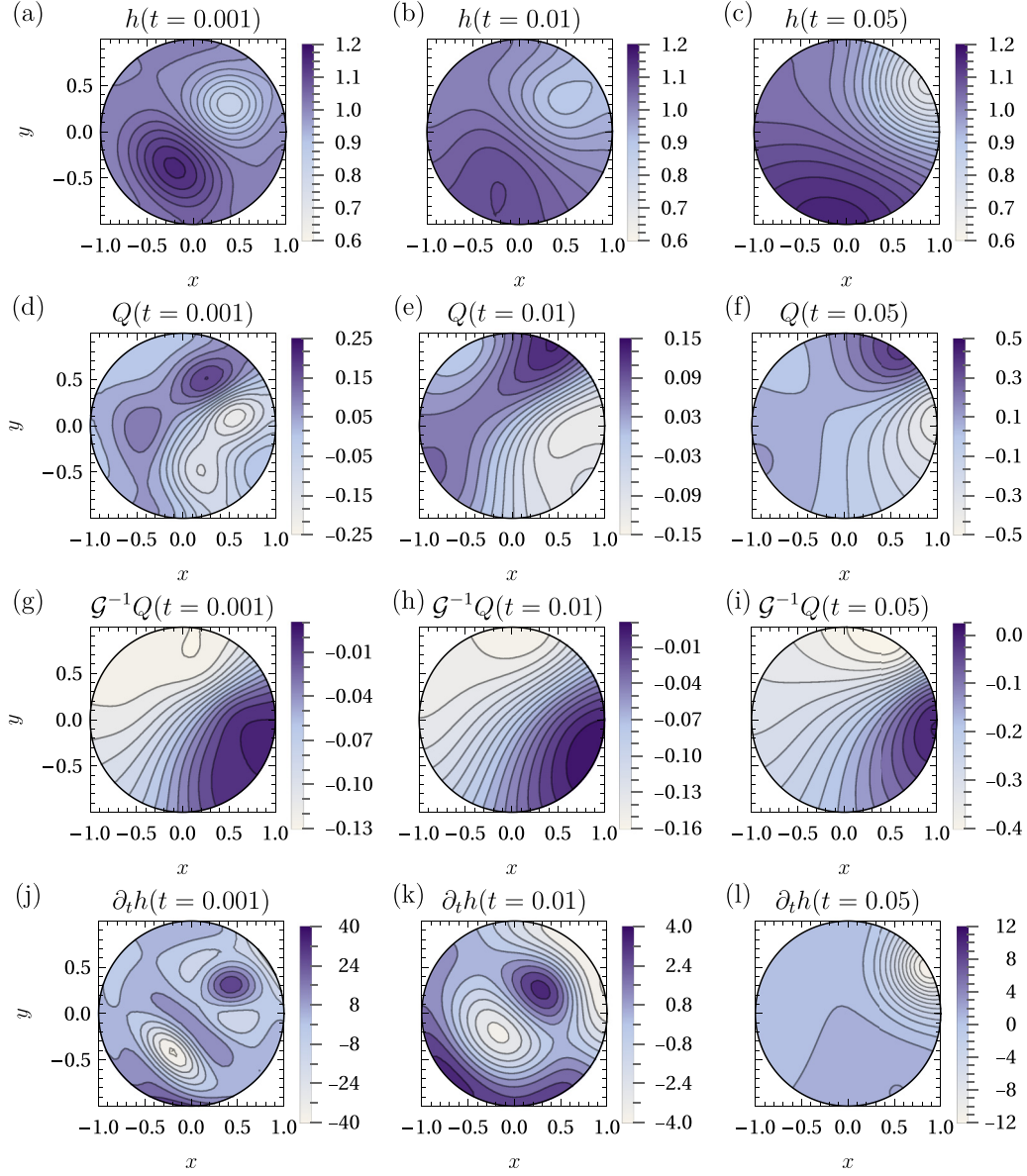


FIG. 2. (a)–(c) Plots of fluid film thickness h , (d)–(f) rotational symmetry characteristic Q , (g)–(i) orthogonal function $\mathcal{G}^{-1}Q$, and (j)–(l) evolution vector $\partial_t h$ at times $t = 0.001, 0.01, 0.05$ during the evolution of the thin-film equation with constant surface tension and van der Waals forces, Eq. (12), for a finite element simulation with mesh element size 0.05 (“medium mesh”). Note that the color scale varies between plots for all quantities except h . As the system evolves, $\mathcal{G}^{-1}Q$ always remains L^2 -orthogonal to $\partial_t h$; equivalently, Q always remains $\mathcal{G}^{-1}$ -orthogonal to $\partial_t h$.

motion is L^2 -orthogonal to the hypothetical pressure driving a rotating film, and the actual pressure is L^2 -orthogonal to the hypothetical film surface motion of a rotating film.

3. Quantifying symmetry violation of numerical simulations

We now use the same system to demonstrate the application of the analytically determined evolutionary constraints to quantify the symmetry violation of numerical simulations. In particular, this provides a new metric for judging the relative quality of simulations: a simulation with only a small violation of a continuous symmetry inherent in the gradient flow equation should be preferred over a simulation with a large violation.

We define the symmetry violation quantity to be $\langle \partial_t h, Q \rangle_{\mathcal{G}^{-1}} = \langle \partial_t h, \mathcal{G}^{-1}Q \rangle_{L^2}$, where Q is again the characteristic of the rotational symmetry. Because this quantity vanishes analytically, in numerical simulations it provides a metric of how well the simulation obeys the governing symmetry. Because $\partial_t h$ and Q vary over time, we also introduce a normalized version of this metric,

$$\frac{\langle \partial_t h, \mathcal{G}^{-1}Q \rangle_{L^2}}{\|\partial_t h\|_{L^2} \|\mathcal{G}^{-1}Q\|_{L^2}}. \quad (23)$$

Results are shown in Fig. 3. The absolute symmetry violation quantity is shown in Fig. 3(a), and the normalized quantity in Fig. 3(b). In both cases, it is clear that the coarse mesh (black circles) has the highest violation, followed by the medium

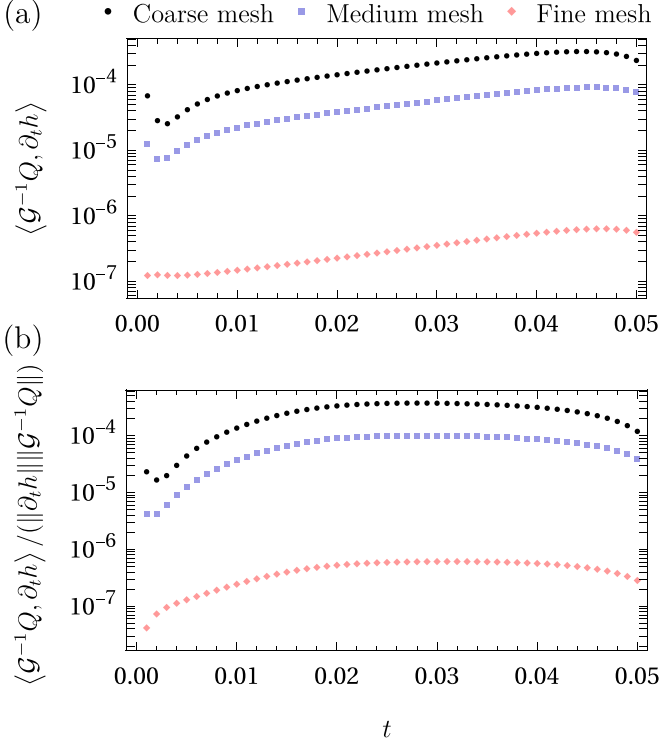


FIG. 3. Comparison of symmetry condition violation between meshes over time: coarse mesh (black circles), with element side length 0.1, medium mesh (blue squares), with element side length 0.05, and fine mesh (pink diamonds), with element side length 0.01. The vertical axis shows the symmetry violation quantity on a log-scale; the horizontal axis is time t . (a) Absolute symmetry violation: L^2 inner product of rotational symmetry characteristic and evolution vector, $\langle \mathcal{G}^{-1}Q, \partial_t h \rangle$. (b) Normalized symmetry violation: L^2 inner product of rotational symmetry characteristic and evolution vector, $\langle \mathcal{G}^{-1}Q, \partial_t h \rangle / (\|\partial_t h\| \|\mathcal{G}^{-1}Q\|)$. Analytically, these values must vanish, but in numerical simulations, they do not, hence quantifying the simulation's violation of the system's symmetry. As the mesh is refined, the violation becomes smaller, indicating a better simulation.

mesh (blue squares), and finally the fine mesh (pink diamonds), which has the lowest violation. Thus, as expected, the finer the mesh, the less the simulation violates the rotational symmetry of the PDE.

These results suggest that the symmetry condition is indeed a useful quantitative metric to assess the quality of a simulation; if the violation is too large, one may wish to use a finer mesh or change the numerical method. It should be noted that both the absolute and normalized symmetry violation quantities vary in time, which is unavoidable due to the changing forms of the functions $\partial_t h$ and Q . Thus, these metrics are most useful as relative comparisons between simulations of the same system at the same point in time, and even the normalized quantity may not be meaningful as an absolute value out of context.

B. Example 2: A gradient flow equation with a conservation law

1. Identification of equation with conservation law

Although the evolutionary constraints of gradient flow problems do not always lead to conserved quantities,

Corollary 1 showed that such cases sometimes occur. Consider the gradient flow equation

$$\langle \partial_t u, w \rangle_{H^{-1}} = \langle \delta_u \mathcal{F}, w \rangle, \quad (24)$$

i.e., $\partial_t u = -\nabla^2 \delta_u \mathcal{F}$, and $\mathcal{G} = -\nabla^2$. We work backward to determine a class of functionals $\mathcal{F}$ such that the gradient flow equation has a conserved quantity associated with a variational symmetry.

Following Corollary 2, a symmetry Q of $\mathcal{F}$ will induce a conserved quantity via $Q = -\nabla^2 \delta_u \mathcal{R}$ if $\mathbf{D}_{\nabla^{-2}} Q = [\mathbf{D}_{\nabla^{-2}} Q]^*$ (although Corollary 2 did not strictly cover the case of $\mathcal{G}^{-1} = -\nabla^{-2}$, we will see that it still holds here). Two immediately obvious values of Q that satisfy this Fréchet derivative condition are $Q = 1$ and $Q = u$. Consider the latter, $Q = u$, resulting in the conserved quantity

$$\mathcal{R} = \int_{\Omega} \mathcal{R}^k dx = -\frac{1}{2} \|u\|_{H^{-1}}^2, \quad (25)$$

$$\mathcal{R} = \frac{1}{2} \nabla s \cdot \nabla s, \quad (26)$$

with s being the solution of $\nabla^2 s = u$ (note that $\delta_u \mathcal{R} = \delta_u [\nabla s \cdot \nabla s / 2] = -s$, and hence $\nabla^2 \delta_u \mathcal{R} = -u$). Boundary conditions on s must be chosen to make boundary terms of the conservation equation vanish, as will be shown.

We thus wish to find a functional $\mathcal{F} = \int \mathcal{F}^k dx$ with $u \delta_u$ as a variational symmetry, i.e., a functional invariant under the scaling $u \rightarrow \lambda u$, for constant λ . There is a wide variety of such functionals, including not only those containing functions of balanced fractions such as $[(\partial_{j_1} u)/u]^{a_1} [(\partial_{j_2} u)/u]^{a_2} \dots$, which explicitly preserve $\mathcal{F}[u] = \mathcal{F}[\lambda u]$, but also those satisfying $\mathcal{F}[u] = \mathcal{F}[\lambda u] + \nabla \cdot \vec{B}[u]$, one example of which is $\mathcal{F} = c_0 \ln(u) + c_1 \ln(\partial_{j_1} u) + \dots$. For this example, let

$$\mathcal{F} = \ln(u), \quad (27)$$

so that

$$\partial_t u = -\nabla^2 \delta_u \mathcal{F} = -\nabla^2 u^{-1}, \quad (28)$$

which is a singular fast-diffusion equation [22].

It is easy to verify by substitution that $\mathcal{R}$ is indeed a conserved quantity of this equation:

$$\begin{aligned} \partial_t \mathcal{R} &= \int_{\Omega} \partial_t \left[\frac{1}{2} \nabla s \cdot \nabla s \right] dx \\ &= \int_{\partial\Omega} s \partial_n (\partial_t s) d^{k-1} x + \int_{\Omega} s \nabla^2 u^{-1} d^k x \\ &= \int_{\partial\Omega} [s \partial_n (\partial_t s + u^{-1}) - u^{-1} \partial_n s] d^{k-1} x + \int_{\Omega} 1 d^k x \\ &= \int_{\partial\Omega} \left[s \partial_n (\partial_t s + u^{-1}) - u^{-1} \partial_n s + \frac{1}{k} \hat{n} \cdot \vec{x} \right] d^{k-1} x, \end{aligned} \quad (29)$$

where $\hat{n}$ represents the normal vector to the boundary and $\partial_n = \hat{n} \cdot \nabla$. Thus $\partial_t \mathcal{R}$ is equal to a boundary integral, which can be controlled by setting the corresponding choice of boundary conditions for s .

The complicated boundary condition, and in particular the presence of $\partial_t s$, makes this result not entirely satisfying. But in the case that u satisfies the Neumann boundary condition $\partial_n u = 0$, then we can set the s boundary condition to $\partial_n s = P^{-1} \int_{\Omega} u d^k x$, where P is the domain perimeter,

$P = \int_{\partial\Omega} 1 d^{k-1}x$ (we cannot, in general, set $\partial_n s = 0$ due to the compatibility condition $\int_{\partial\Omega} \partial_n s d^{k-1}x = \int_{\Omega} u d^k x$). This yields

$$\begin{aligned} \partial_t \mathcal{R} &= \int_{\Omega} 1 d^k x - P^{-1} \int_{\partial\Omega} u^{-1} d^{k-1}x \int_{\Omega} u d^k x \\ &= V - P^{-1} \left[\int_{\partial\Omega} u^{-1} d^{k-1}x \right] \left[\int_{\Omega} u(t=0) d^k x \right], \end{aligned} \quad (30)$$

where V is the domain size, $V = \int_{\Omega} 1 d^k x$, and we have emphasized that the volume integral of u may be taken at $t = 0$, because u 's volume is conserved. This expression still depends on a boundary integral of u^{-1} , but in cases where the domain is large, and the variation in u is confined to a smaller region within the domain, that boundary integral term may be approximated as constant. In particular, if the far-field value of u is a constant u_c , and the mean value of $u(t=0)$ is u_m , then $\partial_t \mathcal{R} \approx V(1 - u_m/u_c)$, i.e.,

$$\mathcal{C}(t) = \mathcal{R}/V - (1 - u_m/u_c)t \quad (31)$$

is approximately constant in time. Note that this is similar to Noether's theorem applied to Lagrangian problems, where quantities such as energy are often constant only with specific boundary conditions, or with fields assumed to become constant in the far field. In the ODE case, e.g., as considered by the aforementioned machine-learning studies [8,9,15], there are no boundary conditions.

A similar argument can be made for other scaling-invariant functionals; for example, $\mathcal{F} = -\ln(\partial_x u) - \ln(\partial_y u)$, corresponding to the fourth-order subdiffusion-like equation $\partial_t u = -\nabla^2[\partial_{xx}u/(\partial_x u)^2 + \partial_{yy}u/(\partial_y u)^2]$. And likewise for other symmetries. For example, the u -translation symmetry ∂_u , corresponding to $Q = 1$, arises when $\delta_u \mathcal{F}$ contains derivatives of u but not u itself, e.g., $\mathcal{F} = (1/2)\nabla u \cdot \nabla u$, corresponding to the biharmonic diffusion equation, $\partial_t u = -\nabla^4 u$. In such cases, Eq. (24) yields the conserved quantity $\int_{\Omega} (x^2 + p)u d^k x$, where p is a harmonic function chosen to satisfy the boundary conditions.

These examples demonstrate that, even with a nontrivial inner product defining the gradient flow equation, there exist large classes of equations with conserved quantities arising due to variational symmetries.

2. Numerical demonstration of conserved quantity

We now demonstrate numerically that the quantity $\mathcal{C}(t)$ defined by Eq. (31) indeed becomes closer to a constant as the domain size is increased, and thus can be used as a metric to determine whether a numerical domain intended to simulate far-field boundary conditions is sufficiently large. Keep in mind that for some gradient flow equations, $\mathcal{R}(t)$ itself may be constant or linear in time regardless of the domain size; in such cases, it could still be used to validate other aspects of the simulation, such as mesh fineness.

We perform simulations in two dimensions using the initial condition given by Eq. (22), which consists of summed Gaussians centered within a distance 0.5 of the origin, and with a maximum length scale of 0.5, so that the nonuniform data are effectively confined to within a few multiples of 0.5 away from the origin.

Finite element simulations of Eq. (28) are performed with time $t \in [0, 0.05]$, time step 0.001, and the backward Euler

method, exactly as in the thin-film equation in Sec. IV A. The mesh has a typical side length of 0.05 ("medium mesh"), and circular domains of four different radii are tested: $r = 1$ (2742 elements), $r = 2$ (11 082 elements), $r = 4$ (43 841 elements), and $r = 8$ (174 654 elements).

Figure 4 shows the evolution of the dependent variable u in a small domain of radius $r = 1$ and a large domain of radius $r = 4$. While the evolution is similar, with the bump and indentation leveling as expected of diffusion, the size of the domain clearly affects the evolution. In particular, the evolution in the small domain interacts with the Neumann boundary condition and produces varying values of u on the boundary, violating the conditions for $\mathcal{C}(t)$ to be constant. In the large domain, there is no visually apparent interaction.

Figure 5(a) shows the quantity $\mathcal{C}(t)$, and its absolute difference from the initial time, $|\mathcal{C}(t) - \mathcal{C}(0)|$, in Fig. 5(b). It is apparent that $\mathcal{C}(t)$ takes different values for different domain sizes. This occurs because $u \approx 1$ in most of the domain; hence $s \approx O(r^2)$, $\mathcal{R} \approx O(r^4)$, and $\mathcal{C} \approx O(r^2)$. Although $\mathcal{C}(t)$ appears constant visually in Figs. 5(a), 5(b) shows that it in fact varies in time, with the scale of the variation decreasing as the domain size increases. Note that the absolute difference is plotted; the relative difference $|\mathcal{C}(t) - \mathcal{C}(0)|/\mathcal{C}(0)$ would show an even more pronounced difference between the simulations.

Knowing that analytically $\mathcal{C}(t)$ must be a constant in a domain with constant far-field boundary conditions, we thus have a metric to assess whether our numerical domain boundary is sufficiently far away. In this example, $\mathcal{C}(t)$ is constant to within a factor of 6×10^{-5} for $r = 1$ (a relatively large value due to the significant boundary effects), 2×10^{-7} for $r = 2$, 6×10^{-8} for $r = 4$, and 2×10^{-8} for $r = 8$, a result which can help guide the choice of domain size analogously to choosing based on energy leakage in a simulation of an energy-conserving Lagrangian system.

V. DISCUSSION

This work has shown how the continuous symmetry characteristics of a gradient flow functional produce a constraint on the evolution vector of that flow, and, in certain cases, also give rise to conserved quantities. This result was applied to the thin-film equation with capillary and van der Waals forces, a highly nonlinear PDE depending on a nontrivial inner product to be put in gradient flow form. The obedience of numerical simulations of this PDE to the required evolutionary constraint was evaluated, producing a metric of the quality of the numerical method. A conserved quantity arising from a continuous symmetry was demonstrated in a singular fast-diffusion equation; to the best of our knowledge, this conserved quantity has not previously been identified.

A number of questions and potential avenues for development remain. While Corollaries 1 and 2 provided sufficient conditions for conserved quantities to exist, it was not proved whether these conditions are necessary. A study similar to Marcotte *et al.*'s [15] work in ODE systems may also be possible, to establish bounds on the number of conserved quantities. The application to physical PDE systems with second-order time derivatives but damping terms breaking the Lagrangian nature is also worth investigating, following the analogous work in ODEs of Gluch and Urbanke [8] and

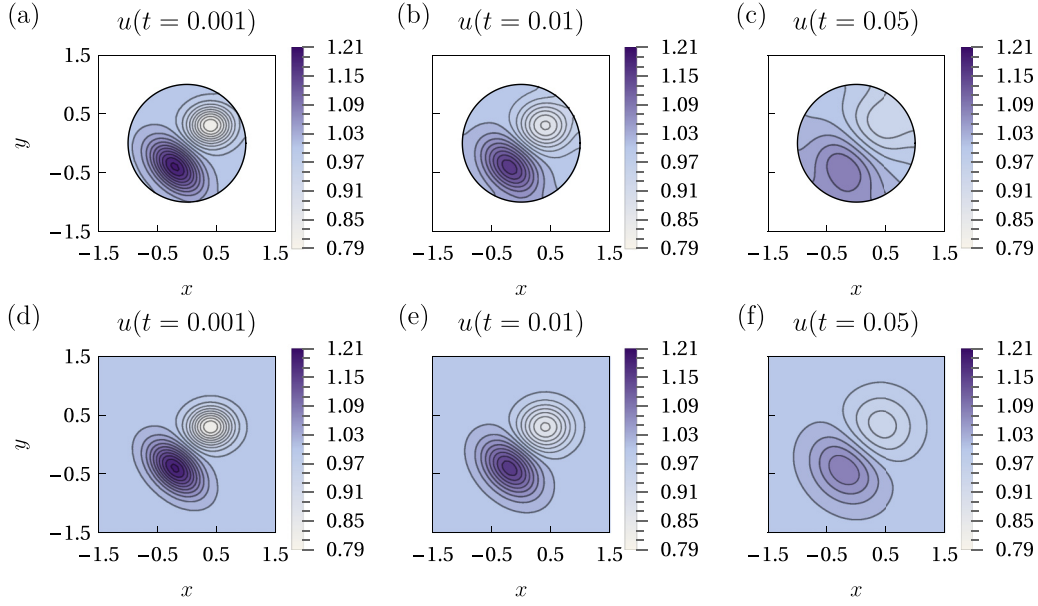


FIG. 4. Plots of the evolution of the dependent variable u of the singular fast-diffusion equation, Eq. (28), at times $t = 0.001, 0.01, 0.05$ for a finite element simulation with mesh element size 0.05 (“medium mesh”). (a)–(c) Small domain (radius $r = 1$), displaying clear boundary effects. (d)–(f) Large domain (radius $r = 4$), where the boundary is sufficiently distant that it has no visually apparent interaction with the evolution.

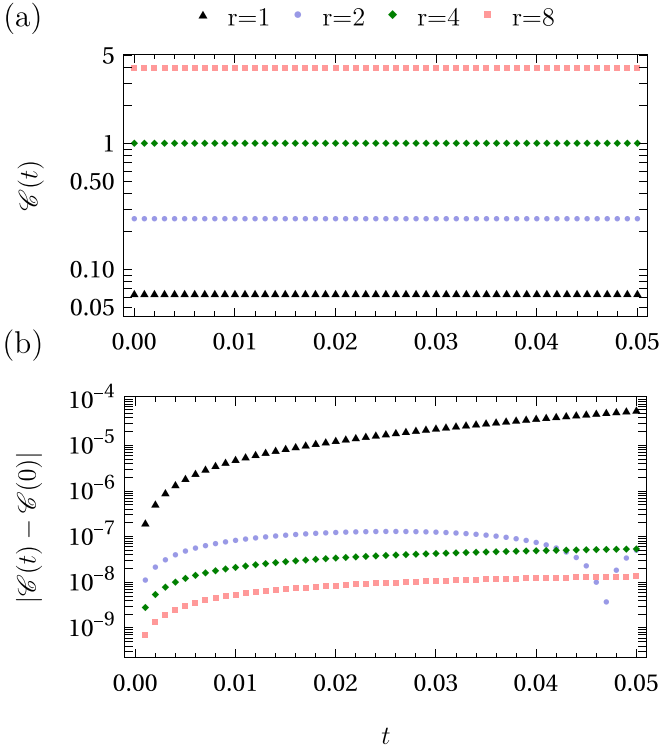


FIG. 5. Comparisons of conserved quantity $\mathcal{E}(t)$, Eq. (31), with u_m and u_c computed at $t = 0$, between domains of radius $r = 1$ (black triangles), $r = 2$ (blue circles), $r = 4$ (green diamonds), and $r = 8$ (pink squares). The horizontal axis is linear-scale and shows time, t ; the vertical axis is log-scale. (a) The quantity itself, $\mathcal{E}(t)$. (b) Absolute change of $\mathcal{E}(t)$ from its initial value. As the domain size is increased, $\mathcal{E}(t)$ becomes closer to constant.

Marcotte *et al.* [16]. Besides the two applications of assessing numerical methods and identifying new conserved quantities, evolutionary constraints may be useful in reducing the dimensionality of or otherwise improving numerical methods. While a single symmetry constrains only one dimension of the evolutionary vector, gradient flow PDEs with infinitely many symmetries (the simplest example being linear PDEs) will have infinitely many constraints, which might lead to more efficient solution methods. It may also be worth pursuing integration methods specifically designed to respect symmetries of the gradient flow, analogous to numerical methods which exactly or nearly conserve energy or momentum in Hamiltonian systems, such as symplectic integrators [23].

More generally, perhaps the most important takeaway of this work should be that Noether’s theorem is not confined to Lagrangian systems or conservation laws. Indeed, any time one works with a variational calculus problem, Noether’s theorem can likely provide some useful insight.

ACKNOWLEDGMENTS

The author would like to thank the anonymous reviewers, whose detailed and insightful feedback substantially improved this work. He would also like to thank Zachary G. Nicolaou, who first introduced him to symmetry analysis. This work is partly based on the author’s Ph.D. thesis [24], which was conducted at the California Institute of Technology (Pasadena, California, USA) and funded by a NASA Space Technology Research Fellowship (80NSSC17K0139).

DATA AVAILABILITY

The simulation code and data that support the findings of this article are openly available [21].

- [1] J. E. Marsden and A. J. Tromba, *Vector Calculus*, 5th ed. (W. H. Freeman and Company, New York, 2003).
- [2] M. I. Jordan, On gradient-based optimization: Accelerated, distributed, asynchronous and stochastic, in *Computational Challenges in Machine Learning* (Simons Institute, Berkeley, California, USA, 2017).
- [3] F. Otto, The geometry of dissipative evolution equations: The porous medium equation, *Commun. Partial Differ. Equ.* **26**, 101 (2001).
- [4] M. M. Weislogel, Capillary flow in interior corners: The infinite column, *Phys. Fluids* **13**, 3101 (2001).
- [5] G. Verbist, D. Weaire, and A. M. Kraynik, The foam drainage equation, *J. Phys.: Condens. Matter* **8**, 3715 (1996).
- [6] L. Giacomelli and F. Otto, Variational formulation for the lubrication approximation of the Hele-Shaw flow, *Calc. Var. PDE* **13**, 377 (2001).
- [7] M. Rumpf and O. Vantzos, Numerical gradient flow discretization of viscous thin films on curved geometries, *Math. Models Methods Appl. Sci.* **23**, 917 (2013).
- [8] G. Gluch and R. Urbanke, Noether: The more things change, the more they stay the same, [arXiv:2104.05508](https://arxiv.org/abs/2104.05508).
- [9] D. Kunin, J. Sagastuy-Brena, S. Ganguli, D. L. K. Yamins, and H. Tanaka, Neural mechanics: Symmetry and broken conservation laws in deep learning dynamics, in *International Conference on Learning Representations* (Vienna, Austria, 2021).
- [10] E. Noether, Invariant variation problems, *Göttingen Nachr.* **1918**, 235 (1918).
- [11] P. J. Olver, On the Hamiltonian structure of evolution equations, *Math. Proc. Cambridge Philos. Soc.* **88**, 71 (1980).
- [12] W. Strampp, On the correspondence between symmetries and conservation laws of evolution equations, *Lett. Math. Phys.* **6**, 113 (1982).
- [13] A.S. Fokas, Generalized symmetries and constants of motion of evolution equations, *Lett. Math. Phys.* **3**, 467 (1979).
- [14] S. S. Du, W. Hu, and J. D. Lee, Algorithmic regularization in learning deep homogeneous models: Layers are automatically balanced, in *Advances in Neural Information Processing Systems*, edited by S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Curran Associates, Inc., 2018), Vol. 31.
- [15] S. Marcotte, R. Gribonval, and G. Peyré, Abide by the law and follow the flow: Conservation laws for gradient flows, in *Advances in Neural Information Processing Systems*, edited by A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, Vol. 36 (Curran Associates, Inc., New Orleans, Louisiana, USA, 2023), pp. 63210–63221.
- [16] S. Marcotte, R. Gribonval, and G. Peyré, Keep the momentum: Conservation laws beyond Euclidean gradient flows, in *41st International Conference on Machine Learning* (Vienna, Austria, 2024).
- [17] P. J. Olver, *Applications of Lie Groups to Differential Equations*, Graduate Texts in Mathematics (Springer-Verlag, New York, 1986).
- [18] P. E. Hydon, *Symmetry Methods for Differential Equations: A Beginner's Guide*, Cambridge Texts in Applied Mathematics (Cambridge University Press, Cambridge, UK, 2000).
- [19] A. Oron, S. H. Davis, and S. G. Bankoff, Long-scale evolution of thin liquid films, *Rev. Mod. Phys.* **69**, 931 (1997).
- [20] F. Hecht, New development in freefem ++, *J. Numer. Math.* **20**, 251 (2012).
- [21] N. C. White, Replication data for beyond Lagrangians: Noether's theorem in gradient flow PDEs, Harvard Dataverse (2025), <https://doi.org/10.7910/DVN/IXTWDB>.
- [22] J. L. Vázquez, *The Porous Medium Equation: Mathematical Theory*, Oxford Science Publications (Clarendon Press, Oxford, UK, 2007).
- [23] J. C. Simo, N. Tarnow, and K. K. Wong, Exact energy-momentum conserving algorithms and symplectic schemes for nonlinear dynamics, *Comput. Methods Appl. Mech. Eng.* **100**, 63 (1992).
- [24] N. C. White, Nonlinear dynamics and stability of viscous free-surface microcapillary flows in V-shaped channels and on curved surfaces, Ph.D. thesis, California Institute of Technology, 2022.