

A Generic Lagrangian-Hamiltonian Framework for Optimal Reference Frames in Unsteady Flow

Xingdi Zhang , Amani Ageeli , Thomas Theußl , Peter Rautek , and Markus Hadwiger 

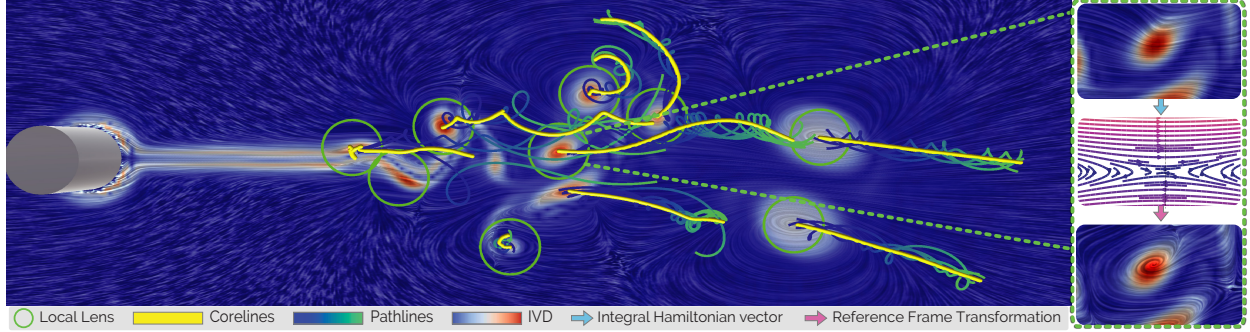


Fig. 1: Optimal observers for different local regions (green circles) in the Boussinesq flow. Each region corresponds to a vortex lens [32]. (Right, middle) Each optimal observer is obtained by integrating a Hamiltonian vector field. (Right, top) Input flow field in the lens region; (right, bottom) flow observed in the lens by the optimal observer. Color coding: IVD. Vortex cores in yellow.

Abstract—Flow visualization inherently depends on a reference frame or observer, relative to which velocities are measured. Recent research has addressed the explicit computation of optimal observers. However, although all techniques share many similarities, the differences can be significant, and no method fulfills all desirable criteria. More accurate methods can be slow, whereas faster approaches can be less accurate, which means that the observers are less optimal. In this paper, we present a novel generic variational framework that combines the best of both worlds, and offers better trade-offs between accuracy and computational efficiency. We build on ideas from Lagrangian and Hamiltonian mechanics to unify previous methods, while at the same time enabling more flexibility and novel algorithmic features. In this way, our framework can also serve as a principled foundation for analysis, comparison, and future advances in this area. We first formulate the observer optimization problem in terms of Lagrangian functions, where an optimal solution must solve the second-order Euler-Lagrange equation. We then expand toward a Hamiltonian perspective, which reduces the second-order equation to first-order. However, we use a hybrid Lagrangian-Hamiltonian approach to combine the advantages of each. By defining Hamiltonian vector fields in an observer phase space, optimal observers simply become integral curves. Our hybrid method computes these vector fields directly from the Lagrangian representing the cost to minimize. Finally, we show how this foundation also enables direct insights into flow field properties, by defining the novel notion of intrinsically steady flows using the Hamiltonian function.

Index Terms—Flow visualization, reference frame optimization, unsteadiness minimization, calculus of variations

1 INTRODUCTION

The computation of optimal reference frames has been established as a powerful ingredient for unsteady flow visualization, and a variety of methods for their optimization have been presented [10, 14, 15, 31–33]. These approaches have many commonalities, but they also have important differences. However, these are often hard to characterize in a mathematically precise way. More accurate approaches can be slow, whereas faster approaches can be less accurate, which means less optimal reference frames. In this paper, we present a novel generic mathematical framework for the variational computation of optimal reference frames using methods from Lagrangian and Hamiltonian mechanics. We combine the best of both worlds: We obtain provably optimal observers in the sense of the Euler-Lagrange equation [1], combined with different options for computational efficiency and desired features, such as progressive computation in time, e.g., for iteratively

computed vortex lenses [32], or optimal for fixed time windows [14]. It is also possible to start with a cheap non-optimal initial guess, and then iterate toward an optimal solution using functional gradient descent.

Our framework offers novel features that are not supported by any previous method, but it also allows unifying, analyzing, and comparing different methods precisely. We first formulate our optimization problem in terms of Lagrangian functions, where an optimal reference frame is a solution of the second-order Euler-Lagrange equation. We then transform the problem from a Lagrangian to a Hamiltonian formulation, which reduces the second-order equation to first-order. Computationally, our method employs a hybrid Lagrangian-Hamiltonian approach: We determine Hamiltonian vector fields and their corresponding optimal trajectories, but use the Lagrangian function which, unlike the Hamiltonian, is directly given by the desired cost function to minimize.

Contributions. The major contributions of our framework are:

- We recast reference frame optimization in a generic way, in a novel unified framework using the calculus of variations [8, 25] in configuration and Lagrangian/Hamiltonian phase spaces [1, 9].
- Our framework supports computing optimal observers using different strategies, guaranteed to solve the Euler-Lagrange equation in the form of Hamilton’s equations and Hamiltonian vector fields.
- Observers can be computed progressively through time simply as integral curves of Hamiltonian vector fields, which avoids drawbacks resulting from fixed time integration windows.

- Xingdi Zhang, Amani Ageeli, Peter Rautek, and Markus Hadwiger are with King Abdullah University of Science and Technology (KAUST), Thuwal, 23955-6900, Saudi Arabia. E-mail: {xingdi.zhang | amani.ageeli | peter.rautek | markus.hadwiger}@kaust.edu.sa
- Thomas Theußl is with Consivi KG, Austria. Email: thomas.theussl@consivi.com

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxx/TVCG.201x.xxxxxx

Table 1: **Method comparison** regarding unsteadiness minimization (d/dt), neighborhood (*region*), integrable in space (*x-integr.*) / time (*t-integr.*), Euclidean observer field (*Killing*) or not (*non-Kill.*), observer guaranteed to match one path line (*path line*), only fixed time interval (*fixed* $[t_a, t_b]$), progressive/incremental computation through time (*t-progress.*), refinement of observer supported (*refine*). Approaches that are not integrable in time (*t-integr.*) do not fulfill the Euler-Lagrange equation for an optimal reference frame. (✓) means each individual optimization step is Euclidean, but due to point-wise computation the resulting reference frame field is in fact not Euclidean. (opt.) means supported but not required.

Method	min. d/dt	region	<i>x-integr.</i>	<i>t-integr.</i>	Killing	non-Kill.	path line	fixed $[t_a, t_b]$	<i>t-progress.</i>	refine
Günther et al. [10]	✓	local	-	-	(✓)	✓	-	-	point-wise t	-
Hadwiger et al. [15]	✓	global	✓	✓	-	✓	-	required	-	-
Kaszás et al. [23]	-	global	✓	✓	✓	-	-	-	point-wise t	-
Rautek et al. [32]	✓	local	✓	-	✓	-	-	-	✓	-
Günther and Theisel [14]	✓	local	/path line	✓	✓	-	✓	required	-	-
Rey Ramirez et al. [33]	✓	local	-	-	(✓)	✓	-	-	point-wise t	-
Our framework	✓	either	✓	✓	✓(opt.)	✓(opt.)	✓(opt.)	✓(opt.)	✓(opt.)	✓(opt.)

- Gradient descent in Hamiltonian phase space allows starting with cheap non-optimal solutions, then converging to the optimum.
- We introduce the novel notion of *intrinsically steady* vector fields, which can be determined directly from their Hamiltonian function.

In future, our framework can serve as a powerful mathematical foundation for designing, analyzing, and comparing observer optimization methods. Our definition of intrinsically steady fields also shows that it can enable determining intrinsic properties of flow fields themselves.

2 RELATED WORK

Extracting and visualizing flow features [30], including vortices [12], is an important area of flow visualization and fluid mechanics.

Reference frame invariance. A crucial aspect is the dependence of flow features on a reference frame or observer. One important type of methods is Galilean invariance [4, 27, 34, 39], which considers the class of inertial reference frames. Several basic properties of flow fields are Galilean-invariant, including spatial derivatives, acceleration [22], vorticity, and feature flow fields [37]. Recently, the more general property of *objectivity* has been considered in continuum mechanics [2, 7, 18, 20, 29, 38] and visualization [11, 15, 36]. Other invariances include rotation invariance [11], similarity and affine invariance [13], and invariance under deformations [3]. Objective flow field properties are the rate-of-strain tensor, spatial derivatives of vorticity, and distances between particles. Several objective feature definitions have been proposed, including the observation of the vorticity tensor in the eigenbasis of the strain rate tensor [2, 7], the strain acceleration tensor [18], instantaneous vorticity deviation (IVD) and its Lagrangian integral (LAVD) [20], and hyperbolic stretching behavior via the finite-time Lyapunov exponent (FTLE) [21, 35]. Objective concepts have been established for all three kinds of Lagrangian coherent structures (LCS) [19] (parabolic, elliptic, and hyperbolic). Reference frames have also been considered in unsteady vector field topology [5].

Reference frame optimization. A recent line of work started by Günther et al. [10] aims to compute optimal reference frames in an objective way, leading to the *objectivity* [36] of properties computed from the corresponding observed flow field. Table 1 categorizes and compares major properties of prior objective methods for computing optimal reference frames. Reference frames can be Euclidean, meaning that their derivative is a Killing vector field [15, 41], or allow some deformation of the frame [3, 13], for example by computing approximate Killing fields [15, 31]. The latter do not directly correspond to a physical Euclidean observer, but individual Euclidean frames can be extracted locally and used globally [40, 41]. A major property that has partially been neglected in previous work is whether reference frames are *integrable* in space and/or time. This aspect is related to whether the number of degrees of freedom an observer computation is using matches the degrees of freedom of the type of reference frame under consideration, as well as to whether the resulting frame obeys the *Euler-Lagrange equation* from the calculus of variations. See Appendix B for a detailed explanation of this property. Günther et al. [10] compute an optimal Euclidean reference frame for each spacetime point (x, t) individually. However, this point-wise computation leads to a reference frame field with more degrees of freedom than is physically possible,

and overall it is *not* a Killing field. Moreover, when the derivatives computed during optimization are integrated in either space or time, they will not match the frames computed at other spacetime points. This property is inherited by the machine learning approach of Kim and Günther [24], and the adaptive approach of Rey Ramirez et al. [33]. The vortex lens of Rautek et al. [32] is integrable in space, but due to simplifications to facilitate progressive computation in time it is not integrable in time. One way of ensuring integrability in time is performing a full least-squares fit in function space, as done by Hadwiger et al. [15] and by Günther and Theisel [14]. The former is integrable in both space and time, but a least-squares fit over all time steps is very slow. In the latter, the frame computed for any individual path line is integrable, but the observer fields from different path lines do not integrate together in space. However, the resulting VDE field, measuring unsteadiness, is a well-defined scalar field. A drawback of both approaches [14, 15] is that they only support fixed integration time windows, and therefore are not suitable for progressive computation through time or vortex lenses [32]. The approach of Kaszás et al. [23] is trivially integrable in space and time, but it does not minimize the unsteadiness of the observed field.

3 BACKGROUND

We build on previous work modeling reference frames as observer fields [15, 31], and Killing vector fields referred to basis vector fields [40, 41]. In the following, we summarize the most important basics.

3.1 Observer Fields

We model reference frame motion using time-dependent observer fields $\mathbf{u}(x, t)$ [15], which are standard velocity fields with special semantics.

Killing fields. For Euclidean observers to preserve distances over time, the corresponding observer field $\mathbf{u}$ must be *Killing* [40, 41], i.e.,

$$\nabla \mathbf{u} + (\nabla \mathbf{u})^T = 0. \quad (1)$$

Because Killing fields form a vector space (and Lie algebra), we can represent any observer as a linear combination of k basis Killing fields [40, 41]: $k = 3$ in 2D ($\mathbb{R}^2$, sphere [31]), and $k = 6$ in 3D ($\mathbb{R}^3$ [40]):

$$\mathbf{u}(x, t) = \sum_{i=1}^k q^i(t) \mathbf{e}_i(x, t), \quad q^i \in \mathbb{R}. \quad (2)$$

Approximate Killing fields. We can loosen the restriction to Euclidean observers by allowing the observer field $\mathbf{u}(x, t)$ to be an *approximate* Killing field [15, 31], where the *Killing energy* $\|\nabla \mathbf{u} + (\nabla \mathbf{u})^T\|_F^2$ (using the Frobenius norm $\|\cdot\|_F$) is minimized but usually not zero. However, we can always extract Euclidean frames locally, e.g., along path lines of the field $\mathbf{u}(x, t)$ as done by Zhang et al. [40, 41]. Moreover, by using spatial basis functions $\mathbf{e}_i$ for general vector fields with k degrees of freedom (e.g., k = number of grid points N · dimensionality n of the domain), we can also represent approximate Killing fields via Eq. 2.

3.2 Observed Vector Field Quantities

Given an observer field $\mathbf{u}(x, t)$, we can define observed vector field quantities. The instantaneous **observed velocity** is simply given

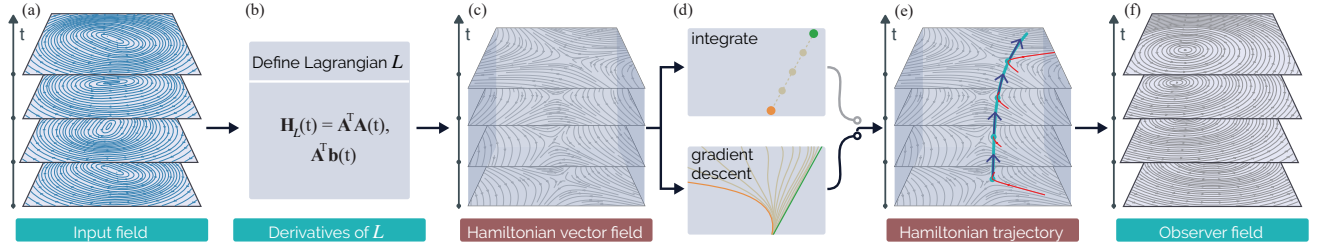


Fig. 2: **Pipeline.** (a) The input field $\mathbf{v}(x, t)$ determines (b) the Lagrangian L and its derivatives, which give (c) the Hamiltonian vector field $\bar{X}_H(\mathbf{q}, \mathbf{p}, t)$ in phase space. (d) Any optimal observer is a particular Hamiltonian trajectory, which we choose by using one of two main strategies: (1) Integrate a Hamiltonian trajectory starting at an initial point in phase space; (2) Use functional gradient descent to converge to an (optimal) trajectory from an initial non-optimal curve. (e, f) The chosen Hamiltonian trajectory directly determines the corresponding observer field $\mathbf{u}(x, t)$.

by $(\mathbf{v} - \mathbf{u})$. Moreover, we can compute the time derivative observed in the frame described by $\mathbf{u}$ as the **observed time derivative** [15, 31]

$$L_{\mathbf{u}}(\mathbf{v} - \mathbf{u}) = \frac{\partial \mathbf{v}}{\partial t} - \frac{\partial \mathbf{u}}{\partial t} + \nabla \mathbf{v}(\mathbf{u}) - \nabla \mathbf{u}(\mathbf{v}). \quad (3)$$

3.3 Lagrangian and Hamiltonian Mechanics

Lagrangian and Hamiltonian mechanics [1, 9, 28] make use of scalar fields called Lagrangians and Hamiltonians in *phase space* in order to determine the evolution of a physical system using the calculus of variations [8]. This evolution is given by an *optimal curve* in phase space: In Lagrangian mechanics, it is the curve that extremizes a functional defined using the Lagrangian; in Hamiltonian mechanics it is an integral curve of the corresponding Hamiltonian vector field. Our method builds heavily on this very general mathematical machinery, but because our goals are different we introduce a hybrid Lagrangian-Hamiltonian approach that is customized for our purposes. Our primary quantity is a scalar cost function represented by the Lagrangian, which measures how well an observer matches the input flow field (how much it makes it appear steady). However, to actually solve for an optimal observer we nevertheless compute integral curves (Hamiltonian trajectories) of the Hamiltonian vector field corresponding to our Lagrangian.

4 PIPELINE

Fig. 2 depicts the main stages of our generic pipeline for computing an optimal observer field $\mathbf{u}(x, t)$ for any unsteady input flow field $\mathbf{v}(x, t)$. Given the input field $\mathbf{v}$, the first two stages (Fig. 2 (b, c)) determine the corresponding *Hamiltonian vector field* $\bar{X}_H(\mathbf{q}, \mathbf{p}, t)$ in the $(2k + 1)$ -dimensional Hamiltonian phase space (including the time dimension), where k is the number of degrees of freedom of an observer per time step. Unlike the typical approach in physics, however, we compute the vector field $\bar{X}_H$ by only using derivatives of the Lagrangian L , instead of the Hamiltonian H . In our context, this is more efficient and it also uses the more meaningful scalar quantity. In physics, often the more interpretable quantity is the Hamiltonian H , because it gives the actual energy of the physical system. In our context, however, the primary quantity is the Lagrangian L , because it is the actual cost that we want to minimize, in particular the magnitude of the observed time derivative (Eq. 3), making the resulting observed field as steady as possible.

However, the hybrid nature of our pipeline is crucial: While the Lagrangian L determines the cost of an observer, *every possible* optimal observer in fact corresponds to a *Hamiltonian trajectory*, i.e., an integral curve of the Hamiltonian vector field $\bar{X}_H$. This also means that *a particular optimal observer must be chosen*: Optimality depends on initial or boundary conditions, and although all trajectories are optimal they in general have different cost. We employ two main strategies (Fig. 2 (d)): (1) Compute a Hamiltonian trajectory from a chosen starting point in phase space; (2) Choose an initial curve that in general is not a Hamiltonian trajectory, and thus not optimal, and then evolve the curve step by step into a Hamiltonian trajectory by using functional gradient descent. Finally, the optimal observer field $\mathbf{u}(x, t)$ (Fig. 2 (f)) is directly determined by the chosen Hamiltonian trajectory (Fig. 2 (e)).

5 A VARIATIONAL FRAMEWORK FOR OBSERVER SPACES

Using Eq. 2, with a given basis $\mathbf{e}_i$, any observer can be described by k time-dependent coefficient functions $q^i(t)$, $i \in \{1, \dots, k\}$. We now

generalize this notion significantly to define *configuration* and *phase spaces* in which observers are represented by smooth curves, and where we can use the calculus of variations to compute optimal observers.

5.1 The Observer Configuration Space

In analogy to the configuration spaces used in Lagrangian mechanics, we define an *observer configuration space* (a manifold), with *generalized coordinates* $\mathbf{q} \in \mathbb{R}^k$. Any observer is then given by a curve

$$t \mapsto \mathbf{q}(t) = (q^1, \dots, q^k)^T(t). \quad (4)$$

See Fig. 3 (a, b). We require the curve $t \mapsto \mathbf{q}(t)$ to be smooth (at least C^2), with the corresponding *tangent vector field* given by the function

$$t \mapsto \dot{\mathbf{q}}(t) = (\dot{q}^1, \dots, \dot{q}^k)^T(t) = \frac{d}{dt}(q^1, \dots, q^k)^T(t). \quad (5)$$

Here, we use the common notation $\dot{q}^i(t) := (d/dt)q^i(t)$. The tangent vectors $\dot{\mathbf{q}}(t)$ determine the partial time derivative of the observer field (Fig. 3 (c)) given by $t \mapsto \mathbf{q}(t)$, because the time derivative of Eq. 2 is

$$\frac{\partial}{\partial t} \mathbf{u}(x, t) = \sum_{i=1}^k \dot{q}^i(t) \mathbf{e}_i(x, t) + \sum_{i=1}^k q^i(t) \frac{\partial}{\partial t} \mathbf{e}_i(x, t). \quad (6)$$

We note that here, unlike previous work [40, 41], we do not always assume that the time derivative of the basis fields $\mathbf{e}_i$ is zero. However, in this paper we use this capability only to be able to incorporate the approach of Günther and Theisel [14]. In all other cases, $\partial \mathbf{e}_i / \partial t = 0$.

5.2 The Lagrangian Observer Phase Space

Our goal is now to go beyond the k -dimensional configuration space, and additionally be able to define a *scalar field* that is allowed to depend on an observer field, its time derivative, and on time. For $\mathbf{q}, \dot{\mathbf{q}} \in \mathbb{R}^k$, and $t \in \mathbb{R}$, we therefore define a scalar function $L(\mathbf{q}, \dot{\mathbf{q}}, t)$ as a map

$$L: \mathbb{R}^k \times \mathbb{R}^k \times \mathbb{R} \rightarrow \mathbb{R}, \quad (\mathbf{q}, \dot{\mathbf{q}}, t) \mapsto L(\mathbf{q}, \dot{\mathbf{q}}, t). \quad (7)$$

We call the space given by $(\mathbf{q}, \dot{\mathbf{q}}, t)$, with generalized coordinates $\mathbf{q}$ and generalized velocities $\dot{\mathbf{q}}$, the *Lagrangian observer phase space*,¹ and the scalar function L the *Lagrangian function*. See Fig. 3 (d).

As in Lagrangian mechanics, where different Lagrangians can be chosen to be suitable for particular contexts, our motivation for the definition of a general function L is similarly that it allows us to unify a variety of criteria whose goal is to quantify how well a particular reference frame motion matches a given input flow field. For each criterion, we can define a corresponding scalar Lagrangian function L .

In Lagrangian phase space, an observer now corresponds to a curve

$$t \mapsto (\mathbf{q}, \dot{\mathbf{q}}, t)(t) \in \mathbb{R}^k \times \mathbb{R}^k \times \mathbb{R}. \quad (8)$$

This phase space has the key advantage over the lower-dimensional observer configuration space $\mathbb{R}^k$ that any scalar function L can depend on $\dot{\mathbf{q}}$ as *independent* variables, instead of on a particular curve $t \mapsto \mathbf{q}(t)$. A field L depending on $\dot{\mathbf{q}}$ cannot be defined in configuration space, because the same coordinate $\mathbf{q}$ would map to multiple scalar values.

¹The space $(\mathbf{q}, \dot{\mathbf{q}})$ is in fact the *tangent bundle* of the configuration space.

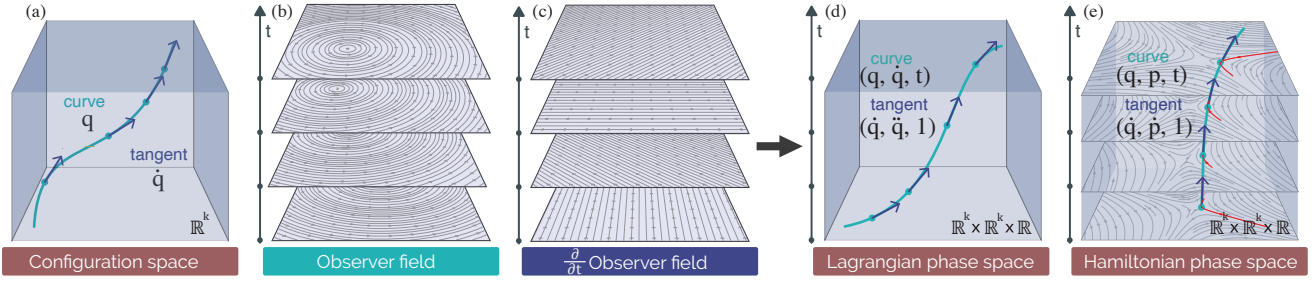


Fig. 3: **Observer spaces.** (a) In *configuration space*, observers are curves $t \mapsto \mathbf{q}(t) \in \mathbb{R}^k$. (b) Each point $\mathbf{q}(t)$ on the curve determines the observer field $\mathbf{u}(\mathbf{x}, t)$ at time t , with curve tangents $\dot{\mathbf{q}}(t)$ giving its time derivative (c). We enable variational observer optimization by going from configuration space to (d) Lagrangian and (e) Hamiltonian *phase spaces*, where observers become curves $t \mapsto (\mathbf{q}, \dot{\mathbf{q}}, t)$, and $t \mapsto (\mathbf{q}, \mathbf{p}, t)$, respectively. Optimality of observers is quantified by a scalar Lagrangian function L in (d), with optimal observers minimizing the integral of L (the *action*). In Hamiltonian phase space, optimal observers are given by trajectories (path lines) of the corresponding *Hamiltonian vector field* (e).

5.3 Objective Observer-Relative Lagrangians

In our framework, the Lagrangian L is a scalar cost function that quantifies how well an observer $\mathbf{q}(t)$ matches an unsteady input flow field $\mathbf{v}(\mathbf{x}, t)$ for each individual time t . We define the function L by integrating a weighted sum of individual cost function terms L_x , defined for spatial points $\mathbf{x}$, over a time-dependent spatial domain $U(t)$,

$$L(\mathbf{q}, \dot{\mathbf{q}}, t) = \frac{1}{2} \int_{U(t)} \left(L_x^D + \lambda_K L_x^K + \lambda_S L_x^S \right) (\mathbf{q}, \dot{\mathbf{q}}, t) d\mathbf{x}. \quad (9)$$

The weights $\lambda_K, \lambda_S \geq 0$ determine the importance of the individual terms, and can be used to simplify the generic cost function for different scenarios by choosing zero weights where desired. In particular, we set $\lambda_K = 0$ for Killing observer fields [14, 32, 40], and only set $\lambda_S > 0$ when additional regularization is desired, as optionally done in prior work [14, 15, 31]. Because our framework is generic, other cost terms L_x can be added in future. However, it is crucial that all terms are *objective* (see Appendix E). Here, we focus on reproducing all cost functions of prior work, and we set the λ weights as done there. We emphasize that our framework nevertheless covers a much wider variety of optimal observer computations, because the cost is only *one ingredient* of determining observers. With detailed derivations in Appendix L, all cost functions of previous work [10, 14, 15, 31–33, 40] map to

$$L_x^D(\mathbf{q}, \dot{\mathbf{q}}, t) = \left\| \dot{q}^i \left(\nabla \mathbf{v}(\mathbf{e}_i) - \nabla \mathbf{e}_i(\mathbf{v}) - \frac{\partial \mathbf{e}_i}{\partial t} \right) - \dot{q}^i \mathbf{e}_i - \mathbf{b} \right\|_{(x, t)}^2, \quad (10a)$$

$$L_x^K(\mathbf{q}, \dot{\mathbf{q}}, t) = \left\| \dot{q}^i \left(\nabla \mathbf{e}_i + (\nabla \mathbf{e}_i)^T \right) \right\|_{(x, t)}^2, \quad (10b)$$

$$L_x^S(\mathbf{q}, \dot{\mathbf{q}}, t) = \left\| \mathbf{s} - \dot{q}^i \mathbf{e}_i \right\|_{(x, t)}^2. \quad (10c)$$

For brevity, here we have used the Einstein summation convention, where upper $(q^i, \dot{q}^i)$ and lower $(\mathbf{e}_i)$ index pairs must be summed over all $i \in [1, k]$ implicitly [16]: $\dot{q}^i \mathbf{e}_i$ must be read as $\sum_{i=1}^k \dot{q}^i \mathbf{e}_i$, etc.

The **observed time derivative** (Eq. 3), relative to an observer field $\mathbf{u}(\mathbf{x}, t)$ given by Eq. 2, is determined by the term L_x^D at a point $\mathbf{x}$. We use this observed time derivative term in three different ways:

- The observer field is a Killing field ($\lambda_K = 0$), with k degrees of freedom ($k = 3$ for 2D flows [41], $k = 6$ for 3D flows [40]). Here, Eq. 10a simplifies for time-independent basis Killing fields $\mathbf{e}_i$, i.e., $\partial \mathbf{e}_i / \partial t = 0$; $\mathbf{b}(\mathbf{x}, t) = -\partial \mathbf{v}(\mathbf{x}, t) / \partial t$ (Appendix L.1.1).
- The observer field is an *approximate* Killing field given by a general vector field $\mathbf{u}(\mathbf{x}, t)$, following Hadwiger et al. [15], and thus $\lambda_K > 0$. Here, the basis functions $\mathbf{e}_i$ are general vector field basis functions; $\mathbf{b}(\mathbf{x}, t) = -\partial \mathbf{v}(\mathbf{x}, t) / \partial t$ (Appendix L.1.2).
- The observer field is Killing ($\lambda_K = 0$), but constrained to contain a path line of the input field $\mathbf{v}(\mathbf{x}, t)$, following Günther and Theisel [14]. Thus, $\mathbf{u}(\mathbf{x}, t)$ does not have full degrees of freedom, requiring time-dependent basis fields $\mathbf{e}_i$ with $\partial \mathbf{e}_i / \partial t \neq 0$ in Eq. 10a, and a specific $\mathbf{b}(\mathbf{x}, t)$. The domain $U(t)$ is centered around the path line, i.e., it is time-dependent. See Appendix L.2.

The **Killing energy** [15] at a point $\mathbf{x}$ is determined by the term L_x^K , and is only required if $\mathbf{u}(\mathbf{x}, t)$ is an approximate Killing field [15, 31].

Seed observer field. The term L_x^S allows requiring that the observer field $\mathbf{u}(\mathbf{x}, t)$ is close to an arbitrary field $\mathbf{s}(\mathbf{x}, t)$. Optional regularization as in prior work [14, 15, 31] results from $\lambda_S > 0$, $\mathbf{s}(\mathbf{x}, t) = \mathbf{v}(\mathbf{x}, t)$. In fact, the Lagrangian then minimizes the *observed kinetic energy* [23].

Quadratic Lagrangians. In our framework, we exploit that Eqs. 9, 10 are quadratic. Specifically, Eq. 9 can be written in the general form

$$L(\mathbf{q}, \dot{\mathbf{q}}, t) = \frac{1}{2} \left\| \mathbf{A}(t) \begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix} - \mathbf{b}(t) \right\|^2. \quad (11)$$

Here, $\mathbf{A}(t)$ is an $m \times 2k$ matrix, and $\mathbf{b}(t)$ is an m -dimensional vector. As above, the number k is the dimensionality of the *configuration space*, and m is a constant that depends on the specific definition of L . See Appendix L for the explicit $\mathbf{A}(t)$, $\mathbf{b}(t)$ that we use. From Eq. 11, we determine partial derivatives with respect to $\mathbf{q}, \dot{\mathbf{q}}$, as given in Appendix C. In particular, we obtain the $2k \times 2k$ Hessian matrix

$$\mathbf{H}_L(t) = \mathbf{A}^T \mathbf{A}(t) = \begin{pmatrix} L_{\mathbf{q}\mathbf{q}} & L_{\mathbf{q}\dot{\mathbf{q}}} \\ L_{\dot{\mathbf{q}}\mathbf{q}} & L_{\dot{\mathbf{q}}\dot{\mathbf{q}}} \end{pmatrix} (t). \quad (12)$$

Because L has the form of Eq. 11, the Hessian $\mathbf{H}_L$ is *independent of the generalized coordinates* $\mathbf{q}, \dot{\mathbf{q}}$; it only depends on the choice of $\mathbf{A}(t)$.

5.4 The Euler-Lagrange Equation

Given a Lagrangian function $L(\mathbf{q}, \dot{\mathbf{q}}, t)$, and an observer defined as a curve $t \mapsto \mathbf{q}(t)$, we define a functional called the “*action*” integral [9]

$$S[\mathbf{q}] = \int_{t_a}^{t_b} L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t) dt. \quad (13)$$

In our framework, the action is the total cost that we want to minimize. In particular, for L_x^D (Eq. 10a), the action measures the total observed time derivative (its squared magnitude) in the time interval $[t_a, t_b]$.

From the calculus of variations, for any curve $t \mapsto \mathbf{q}(t)$, $t \in [t_a, t_b]$, minimizing² the action $S[\mathbf{q}]$ in Eq. 13, we must fulfill the corresponding *Euler-Lagrange equation* [8]. See Appendix J.1. We define the slightly more general vector-valued *Euler-Lagrange operator* ψ , which is

$$\psi = \frac{\partial L}{\partial \mathbf{q}} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{q}}}. \quad (14)$$

Using this operator, the Euler-Lagrange equation is simply $\psi = 0$. If a curve $t \mapsto \mathbf{q}(t)$ fulfills $\psi = 0$ for all $t \in (t_a, t_b)$ in a given time domain, and it is unique (with appropriate initial or boundary conditions), we call it the *optimal observer* for the whole time interval $[t_a, t_b]$. Appendix D gives the system of second-order ODEs that $\psi = 0$ represents, by explicitly expanding the Euler-Lagrange operator ψ in Eq. 14 in terms of first and second partial derivatives of the Lagrangian L .

²Generally, this holds for *stationary points* (in function space). However, in our framework, because of our Lagrangians (Eq. 11) we always have minima.

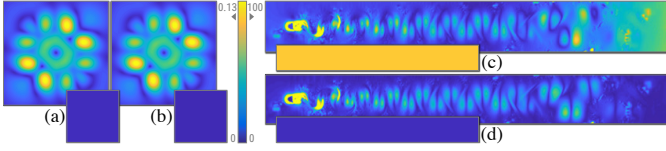


Fig. 4: **The Euler-Lagrange operator** must vanish ($\psi = 0$) for optimal observers. The RFC-C (Four Centers) flow (a,b) is *intrinsically steady*: (a) $\psi = 0$ even for point-wise computation [10]. For the more complicated Cylinder2D flow (c,d), however, (c) $\psi \neq 0$ [10]. Our framework *guarantees* optimal observers with $\psi = 0$ (b,d). Images: color-coded $\|\psi\|$ of per-pixel L_x ; insets: integrated Lagrangian L .

6 THE HAMILTONIAN OBSERVER PHASE SPACE

Instead of solving the second-order ODE $\psi = 0$, we will solve the corresponding first-order *Hamilton's equations* [9] in Hamiltonian phase space, which is equivalent to solving the Euler-Lagrange equation.

As in Hamiltonian mechanics, corresponding to our Lagrangian L we define the k -dimensional covariant³ *generalized momentum*

$$\mathbf{p} := \frac{\partial L}{\partial \dot{\mathbf{q}}}. \quad (15)$$

To prepare for the Hamiltonian function H , for a Legendre transformation [1] of L we define the energy function $E_L(\mathbf{q}, \dot{\mathbf{q}}, t) = \mathbf{p}^T \dot{\mathbf{q}} - L$, and define the *Hamiltonian phase space* $(\mathbf{q}, \mathbf{p}, t) \in \mathbb{R}^k \times \mathbb{R}^k \times \mathbb{R}$ via a transformation of the phase space coordinates $(\mathbf{q}, \dot{\mathbf{q}}, t)$ to coordinates $(\mathbf{q}, \mathbf{p}, t)$. We require that this coordinate transformation is invertible:

$$(\mathbf{q}, \mathbf{p}, t) = (\mathbf{q}, \mathbf{p}(\mathbf{q}, \dot{\mathbf{q}}, t), t), \quad \text{with inverse} \quad (16a)$$

$$(\mathbf{q}, \dot{\mathbf{q}}, t) = (\mathbf{q}, \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t), t). \quad (16b)$$

Given the Lagrangian $L(\mathbf{q}, \dot{\mathbf{q}}, t)$, the Hamiltonian function $H(\mathbf{q}, \mathbf{p}, t)$ in the Hamiltonian phase space is determined by the energy E_L together with the coordinate transformation given by Eq. 16b. That is, we define

$$H(\mathbf{q}, \mathbf{p}, t) = E_L(\mathbf{q}, \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t), t). \quad (17)$$

From the Legendre transformation giving $H(\mathbf{q}, \mathbf{p}, t)$ from $L(\mathbf{q}, \dot{\mathbf{q}}, t)$, we can derive *Hamilton's equations* (see Appendix F). Crucially, we exploit that these equations are not only valid for particular curves $t \mapsto (\mathbf{q}, \mathbf{p}, t)(t)$, but they define a *Hamiltonian vector field* in the phase space $(\mathbf{q}, \mathbf{p}, t)$. To emphasize this, we write Hamilton's equations as

$$\begin{aligned} \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t) &= \frac{\partial H}{\partial \mathbf{p}}(\mathbf{q}, \mathbf{p}, t), \\ \dot{\mathbf{p}}(\mathbf{q}, \mathbf{p}, t) &= -\frac{\partial H}{\partial \mathbf{q}}(\mathbf{q}, \mathbf{p}, t). \end{aligned} \quad (18)$$

In order to determine an optimal observer given by an optimal curve $t \mapsto (\mathbf{q}, \mathbf{p}, t)(t)$ in Hamiltonian phase space, we therefore can now solve a system of $2k$ first-order ODEs, instead of solving the system of k second-order ODEs comprising the Euler-Lagrange equation.

If $\partial H / \partial t = 0$, or if we fix one specific time $t = \tau$, we can define a steady Hamiltonian $2k$ -vector field [9] as $X_H := (\dot{\mathbf{q}}, \dot{\mathbf{p}})^T$. For unsteady flow fields $\mathbf{v}(x, t)$, however, we have $\partial H / \partial t \neq 0$, and we also define a $(2k + 1)$ -vector field $\bar{X}_H(t)$. We define the two fields

$$X_H(\mathbf{q}, \mathbf{p}, \tau) = \begin{pmatrix} \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, \tau) \\ \dot{\mathbf{p}}(\mathbf{q}, \mathbf{p}, \tau) \end{pmatrix}, \quad \bar{X}_H(\mathbf{q}, \mathbf{p}, t) = \begin{pmatrix} \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t) \\ \dot{\mathbf{p}}(\mathbf{q}, \mathbf{p}, t) \\ 1 \end{pmatrix}. \quad (19)$$

A complete definition of Hamiltonian vector fields and their properties is given in Appendix F. However, the crucial point is that we can now exploit the fundamental result in Hamiltonian mechanics that, by Hamilton's equations, *the integral curves of the vector field $\bar{X}_H$ (the Hamiltonian trajectories), are solutions to the Euler-Lagrange equation.*

³It is crucial here that $\partial L / \partial \dot{\mathbf{q}}$ is in fact a tensor (a covector). The Hamiltonian phase space $(\mathbf{q}, \mathbf{p})$ thus comprises the *cotangent bundle* of the configuration space. See Appendix M. (Somewhat surprisingly, $\partial L / \partial \mathbf{q}$ is *not* tensorial.)

7 LAGRANGIAN-HAMILTONIAN METHOD

We will now derive how the general structure of a Lagrangian L in the form of Eq. 11 can be exploited to obtain a form of Hamilton's equations that can be computed without direct computation of the partial derivatives of the Hamiltonian H in Eq. 18. Instead, we can reformulate Hamilton's equations directly using only the Lagrangian, as given by Eqs. 25a, 25b below, which only require the partial derivatives in *Lagrangian phase space* comprising the Hessian matrix in Eq. 12.

The crucial point is that although Eqs. 25a, 25b are *a form of Hamilton's equations*, and therefore they of course give the derivatives in Eq. 18, we do not compute them by partial derivation of H .

7.1 Lagrangian-Hamiltonian Formulation

For a fixed time t , we can set the first-order partial derivatives of the Lagrangian L to zero, to solve for the coordinates $(\mathbf{q}, \dot{\mathbf{q}})$ where $L(\cdot, \cdot, t)$ attains a minimum. However, we can also pre-scribe *non-zero* derivatives $L_{\mathbf{q}} := \partial L / \partial \mathbf{q}$, $\mathbf{p} := \partial L / \partial \dot{\mathbf{q}}$, and solve for $(\mathbf{q}, \dot{\mathbf{q}})$ such that

$$\mathbf{H}_L(t) \begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix} - \mathbf{A}^T \mathbf{b}(t) = \begin{pmatrix} L_{\mathbf{q}} \\ \mathbf{p} \end{pmatrix}. \quad (20)$$

See also Eq. 61. In many equations, we will also need the *upper* and *lower* halves, respectively, of the $2k$ -vector $\mathbf{A}^T \mathbf{b}(t)$, and we define

$$\begin{aligned} \bar{\mathbf{b}}_h(t) &= \mathbf{P}_h \mathbf{A}^T \mathbf{b}(t), \\ \bar{\mathbf{b}}_l(t) &= \mathbf{P}_l \mathbf{A}^T \mathbf{b}(t). \end{aligned} \quad (21)$$

The operators $\mathbf{P}_{\{h|l\}} : \mathbb{R}^{2k} \rightarrow \mathbb{R}^k$ project to the k -dimensional *upper* ($\mathbf{P}_h$) or *lower* ($\mathbf{P}_l$) component half of a $2k$ -vector, respectively.

Using the upper half of the matrix $\mathbf{H}_L$, i.e., $\mathbf{P}_h \mathbf{H}_L$, Eq. 20 gives the partial derivatives $L_{\mathbf{q}}$ at a given coordinate $(\mathbf{q}, \dot{\mathbf{q}}, t)$ as

$$L_{\mathbf{q}}(\mathbf{q}, \dot{\mathbf{q}}, t) = L_{\mathbf{q}\mathbf{q}}(\mathbf{q}) \mathbf{q} + L_{\mathbf{q}\dot{\mathbf{q}}}(t) \dot{\mathbf{q}} - \bar{\mathbf{b}}_h(t). \quad (22)$$

Moreover, using the lower half of the matrix $\mathbf{H}_L$, i.e., $\mathbf{P}_l \mathbf{H}_L$, Eq. 20 gives the *coordinate transformation* $\mathbf{p}(\mathbf{q}, \dot{\mathbf{q}}, t)$, with $(\mathbf{q}, \dot{\mathbf{q}}, t)$ known,

$$\mathbf{p}(\mathbf{q}, \dot{\mathbf{q}}, t) = L_{\dot{\mathbf{q}}\mathbf{q}}(\mathbf{q}) \mathbf{q} + L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}(t) \dot{\mathbf{q}} - \bar{\mathbf{b}}_l(t). \quad (23)$$

In particular, we also need the *inverse* of Eq. 23, i.e., the coordinate transformation $\dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t)$, with $(\mathbf{q}, \mathbf{p}, t)$ known. Requiring that $L_{\dot{\mathbf{q}}\mathbf{q}}$ is invertible (which is fulfilled), this coordinate transformation is given by

$$\dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t) = L_{\dot{\mathbf{q}}\mathbf{q}}^{-1}(t) (\mathbf{p} - L_{\dot{\mathbf{q}}\mathbf{q}}(t) \mathbf{q} + \bar{\mathbf{b}}_l(t)). \quad (24)$$

The crucial point here is this one: If the Lagrangian L is in the form of Eq. 11, it is possible to compute the Hessian matrices required in Eq. 24 *although $\dot{\mathbf{q}}$ is not known yet*, because they are *independent* of $\dot{\mathbf{q}}$.

Now, however, because Eq. 24 solves for $\dot{\mathbf{q}}$, we can compute the term $-\partial H / \partial \mathbf{q} = \partial L / \partial \mathbf{q}$ (cf. Eq. 73) in Eq. 18 directly in Lagrangian instead of in Hamiltonian phase space coordinates. This means that for all Lagrangians L in the form of Eq. 11, by additionally using Eq. 22 we can fully re-write Hamilton's equations using the Hessian $\mathbf{H}_L(t)$ from Eq. 12. This gives *Hamilton's equations* in the form (cf. Eq. 18)

$$\dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t) = L_{\dot{\mathbf{q}}\mathbf{q}}^{-1}(\mathbf{p} - L_{\dot{\mathbf{q}}\mathbf{q}} \mathbf{q} + \bar{\mathbf{b}}_l)(t), \quad (25a)$$

$$\dot{\mathbf{p}}(\mathbf{q}, \mathbf{p}, t) = (L_{\mathbf{q}\mathbf{q}} \mathbf{q} + L_{\mathbf{q}\dot{\mathbf{q}}} \dot{\mathbf{q}} - \bar{\mathbf{b}}_h)(t). \quad (25b)$$

7.2 The Hamiltonian Vector Field

If we remove the dependence of Eq. 25b on $\dot{\mathbf{q}}$ on the right-hand side by inserting Eq. 25a, we furthermore obtain equations depending only on $(\mathbf{q}, \mathbf{p})$. See Appendix F.3. We can thus write Hamilton's equations as a linear *Hamiltonian vector field* $X_H(\mathbf{q}, \mathbf{p}, t)$, given by the affine map

$$\begin{pmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{pmatrix}(\mathbf{q}, \mathbf{p}, t) = X_H(\mathbf{q}, \mathbf{p}, t) = \mathbf{X}(t) \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} + X_H(0, 0, t). \quad (26)$$

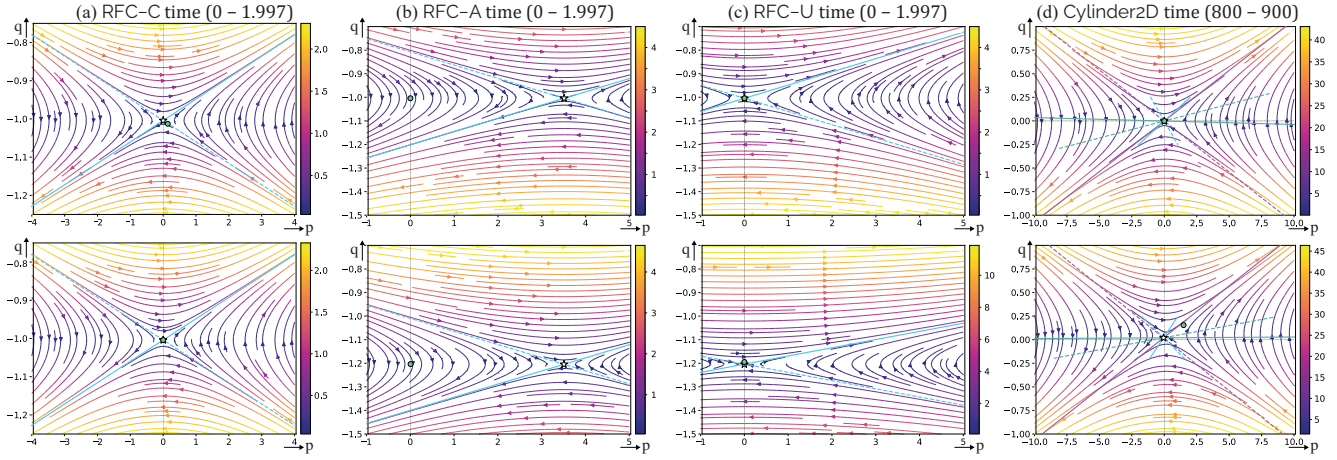


Fig. 5: **Hamiltonian vector fields** for Euclidean 2D observers ($k = 3$). (a) RFC-C, (b) RFC-A, (c) RFC-U, (d) Cylinder2D. Rows show two different time steps, for 2D subspaces of the 6D ($2k$) Hamiltonian phase space; optimal observer $\mathbf{q}(t)$ shown in green; 6D eigenvectors of $\mathbf{X}(t)$ shown projected to 2D. (a) The optimal observer stays on the critical point of $X_H(t)$, because the angular velocity of the observer is constant. (b) The angular velocity q^3 of the optimal observer follows the downward motion of the critical point at a distance. (c) Although this data set is intrinsically unsteady, the optimal observer is the same as in (b). (d) Although it starts at $\mathbf{p} = 0$, the optimal observer leaves the $\mathbf{p} = 0$ subspace.

Importantly, in our Lagrangian-Hamiltonian framework, we determine the field X_H from the Lagrangian L without knowing the Hamiltonian H . The field X_H is affine because our Lagrangians L are quadratic. As derived in Appendix F.3, we obtain the $2k \times 2k$ matrix $\mathbf{X}(t)$ as

$$\mathbf{X}(t) = \begin{pmatrix} -L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}^{-1} L_{\dot{\mathbf{q}}\mathbf{q}} & L_{\dot{\mathbf{q}}\mathbf{q}}^{-1} \\ L_{\mathbf{q}\mathbf{q}} - L_{\mathbf{q}\dot{\mathbf{q}}} L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}^{-1} L_{\dot{\mathbf{q}}\mathbf{q}} & L_{\mathbf{q}\dot{\mathbf{q}}} L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}^{-1} \end{pmatrix} (t). \quad (27)$$

The affine part of Eq. 26 is given by the fixed vector (for each time t)

$$X_H(0, 0, t) = \begin{pmatrix} L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}^{-1} \bar{\mathbf{b}}_l \\ L_{\mathbf{q}\dot{\mathbf{q}}} L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}^{-1} \bar{\mathbf{b}}_l - \bar{\mathbf{b}}_h \end{pmatrix} (t). \quad (28)$$

We can notice that the matrix $\mathbf{X}(t)$ given by Eq. 27 has the following block structure, comprising individual $k \times k$ blocks $\mathbf{A}, \mathbf{B}, \mathbf{C}$. This corresponds to the fact that $\mathbf{X}(t)$ is a *Hamiltonian matrix*, in the form

$$\mathbf{X}(t) = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & -\mathbf{A}^T \end{pmatrix} (t). \quad (29)$$

From Eq. 27, we note the block symmetries $\mathbf{B} = \mathbf{B}^T$, and $\mathbf{C} = \mathbf{C}^T$. Considering the *canonical symplectic matrix* Ω (Appendix F.1) gives

$$\Omega = \begin{pmatrix} \mathbf{0} & -\mathbf{I}^{k \times k} \\ \mathbf{I}^{k \times k} & \mathbf{0} \end{pmatrix}, \quad \Omega \mathbf{X}(t) = \begin{pmatrix} -\mathbf{C} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{B} \end{pmatrix} (t). \quad (30)$$

We note that $\Omega \mathbf{X}$ is symmetric. Moreover, because $\mathbf{X}^T \Omega + \Omega \mathbf{X} = 0$, $\mathbf{X}$ is a Hamiltonian matrix [9]. This also means that $\mathbf{X}$ is an *infinitesimal symplectic matrix*, in the Lie algebra of the symplectic (Lie) group. Hamiltonian vector fields can be integrated as any other vector field, but for improved accuracy we use *symplectic integrators* (Appendix F.6).

7.3 The Hamiltonian

As above, a crucial advantage of our Lagrangian-Hamiltonian approach is that it *does not require the Hamiltonian H explicitly* in order to solve Hamilton's equations in the form of Eq. 25a, 25b, or in the form of Eq. 26 that explicitly gives the Hamiltonian vector field X_H . We nevertheless now derive the general form of the function H when the Lagrangian L is given in the general form of Eq. 11. This is not only useful for completeness, but for analyzing interesting properties such as separability or the time derivative of the Hamiltonian function H .

Interestingly, using our Lagrangian-Hamiltonian approach, we can determine the full equation of the Hamiltonian $H(\mathbf{q}, \mathbf{p}, t)$ by working “backwards” from the Hamiltonian vector field X_H that we already know from Eq. 26. This again saves us tedious computation of partial derivatives of the Hamiltonian H instead of the Lagrangian L .

We first employ a relationship known in Hamiltonian mechanics, which says that the Hessian matrix $\mathbf{H}_H(t)$ of the Hamiltonian $H(\mathbf{q}, \mathbf{p}, t)$ defines the Hamiltonian matrix $\mathbf{X}(t)$ in Eq. 26 by the relationship $\mathbf{X}(t) = -\Omega \mathbf{H}_H(t)$ (with the sign convention of Ω in Eq. 30). From this, without having an expression for the Hamiltonian function $H(\mathbf{q}, \mathbf{p}, t)$, we directly obtain its Hessian matrix (cf. Eq. 30)

$$\mathbf{H}_H(t) = \Omega \mathbf{X}(t). \quad (31)$$

We then just need to compute the Hamiltonian function H in a not fully expanded form (cf. Eq. 111) that allows us to compare with the terms given by the known matrix $\mathbf{H}_H(t)$. We derive the necessary details in Appendix G. In the end, we obtain the Hamiltonian $H(\mathbf{q}, \mathbf{p}, t)$ as

$$H(\mathbf{q}, \mathbf{p}, t) = \frac{1}{2} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix}^T \mathbf{H}_H(t) \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} + \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix}^T \begin{pmatrix} -\dot{\mathbf{p}} \\ \dot{\mathbf{q}} \end{pmatrix} (0, 0, t) + \bar{\mathbf{b}}_l^T L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}^{-1} \bar{\mathbf{b}}_l(t) - \frac{1}{2} \mathbf{b}^T \mathbf{b}(t). \quad (32)$$

The vectors $(\dot{\mathbf{q}}, \dot{\mathbf{p}})(0, 0, t)$ come from $X_H(0, 0, t)$. See Eqs. 26, 28.

7.4 The Euler-Lagrange Operator

The Euler-Lagrange operator $\psi(t)$ is important in our framework, because it gives the *functional gradient* (Sec. 8.3). Using the matrix $\mathbf{X}(t)$ and the vector $\dot{\mathbf{p}}(0, 0, t)$ introduced in Eq. 26 (cf. Eq. 28), we compute

$$\psi(t) = \begin{pmatrix} \mathbf{P}_l \mathbf{X}(t) \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} (t) + \dot{\mathbf{p}}(0, 0, t) - \dot{\mathbf{p}}(t). \quad (33)$$

The matrix $\mathbf{P}_l \mathbf{X}(t)$ is the lower k -half of the matrix $\mathbf{X}(t)$ given by Eq. 27. Because Eq. 33 is the same as $\psi(t) = \dot{\mathbf{p}} - \dot{\mathbf{p}}(t)$, it is crucial to realize that it does *not* say $\psi(t) = \dot{\mathbf{p}} - \dot{\mathbf{p}}(t) = 0$. The reason for this is that $\dot{\mathbf{p}}(t)$ is half of the *actual* tangent of *some* curve $t \mapsto (\mathbf{q}, \mathbf{p}, t)(t)$, whereas $\dot{\mathbf{p}}$ is the vector that Eqs. 26, 27, 28 (or Eq. 25b) say the tangent *has to be* in order to fulfill the Euler-Lagrange equation $\psi(t) = 0$.

8 COMPUTING OPTIMAL OBSERVERS

We consider two major types of initializations for optimal reference frame computation: (1) *Initial value* $(\mathbf{q}, \mathbf{p}, \tau)$ for some arbitrarily chosen time $t = \tau$, from which we compute an integral curve of the Hamiltonian vector field, i.e., a *Hamiltonian trajectory*. (2) *Initial guess* for an entire curve $t \mapsto (\mathbf{q}, \mathbf{p}, t)(t)$, often not fulfilling Hamilton's equations, which is then evolved to an optimal curve via functional gradient descent (Sec. 8.3). Here, we solve a *boundary value problem* with a *free boundary* (natural boundary conditions; see Appendix K).

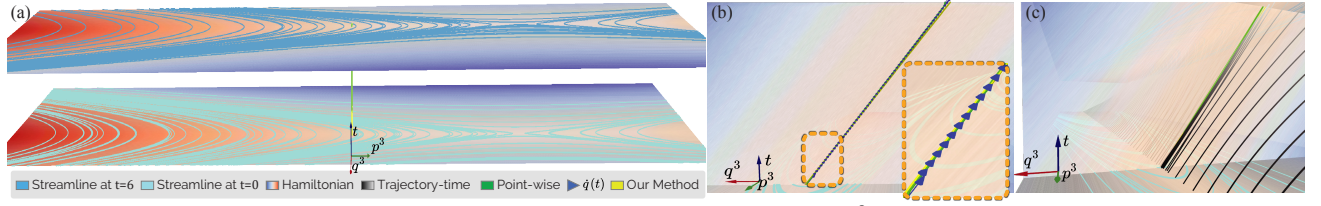


Fig. 6: **Hamiltonian vector field and trajectories (RFC-A flow, $k = 3$).** We depict the 3D (q^3, p_3, t) subspace (time vertical), color-coded scalar Hamiltonian function H (Eq. 32). (Yellow) Optimal observer from Hamiltonian trajectory. (Green) Curve computed point-wise. Both curves agree exactly because this field is intrinsically steady. (a) Hamiltonian stream lines for two time steps. (b) Hamiltonian trajectory of optimal observer in the $\mathbf{p} = 0$ subspace (vertical plane) for all time. (c) Other Hamiltonian trajectories from different initial values ($\mathbf{q} = (0, 0 - 1)$, $\mathbf{p} \neq 0$).

Common setup. If we set the partial derivatives $L_{\mathbf{q}}, \mathbf{p} = \partial L / \partial \dot{\mathbf{q}}$ in Eq. 20 to zero, assuming $\mathbf{H}_L(t)$ is invertible for time t we can solve

$$\mathbf{q} = \mathbf{P}_h \mathbf{H}_L^{-1} \mathbf{A}^T \mathbf{b}(t). \quad (34)$$

This equation computes the generalized coordinate $\mathbf{q}$ of a point $(\mathbf{q}, \dot{\mathbf{q}})$ in Lagrangian phase space where L attains its unique minimum. The corresponding point in Hamiltonian phase space is $(\mathbf{q}, 0)$, i.e., $\mathbf{p} = 0$.

For approximate Killing fields $\mathbf{u}$, where the dimensionality $2k$ is large, instead of solving Eq. 34 via direct matrix inversion, we solve the linear system $\mathbf{H}_L(\mathbf{q}, \dot{\mathbf{q}})^T = \mathbf{A}^T \mathbf{b}(t)$ to obtain $\mathbf{q} = \mathbf{P}_h(\mathbf{q}, \dot{\mathbf{q}})^T$.

8.1 Initial Value for a Hamiltonian Trajectory

We obtain the coordinate $\mathbf{q}$ for a time $t = \tau$ from Eq. 34, and compute a Hamiltonian trajectory starting at the point $(\mathbf{q}, 0, \tau)$ by integrating $\bar{X}_H$, given by Eqs. 19, 26, with $(\mathbf{q}, \mathbf{p}, \tau)$ as initial value (see also Eq. 92).

A crucial caveat here is that in variational problems it is, in general, *incorrect to define a curve $t \mapsto (\mathbf{q}, \mathbf{p}, t)(t)$ point-wise by connecting individual extremal points (here, minima).*⁴ The underlying reason is subtle: First, a curve obtained point-wise will in general not be an integral curve of the Hamiltonian vector field $\bar{X}_H$, and thus it will not solve the Euler-Lagrange equation. Moreover, because a curve defined point-wise by $\mathbf{q}$ from Eq. 34, with $\mathbf{p} = 0$, connects minima, *unless we have the special case that it is a Hamiltonian trajectory*, we even get the more severe problem that its coordinates do not match its tangents:

$$\dot{\mathbf{q}}(\mathbf{q}(t), 0, t) \neq \frac{d}{dt} \mathbf{q}(t). \quad (35)$$

This is the same as saying that the coordinate transformation $\dot{\mathbf{q}}(\mathbf{q}, 0, t)$ using Eq. 24 does not actually give the tangent vector field of the curve, which means that the coordinates $\mathbf{p}(t) = \partial L / \partial \dot{\mathbf{q}} = 0$ do not match the curve. Of course, not using a $\mathbf{p}$ coordinate explicitly and defining a curve $t \mapsto (\mathbf{q}, \dot{\mathbf{q}})^T(t) = \mathbf{H}_L^{-1} \mathbf{A}^T \mathbf{b}(t)$ is equally incorrect. See Appendix H for more details and proofs of these properties. We note that one *special case* for which a point-wise curve is indeed correct are *intrinsically steady* vector fields $\mathbf{v}(x, t)$, as introduced in Sec. 9.

8.2 Initial Non-Optimal Curve in Hamiltonian Phase Space

In contrast to determining just one initial point, for an entire curve we employ the following strategy. We first compute a curve $t \mapsto \mathbf{q}(t)$ using Eq. 34 for each time t , i.e., the curve $t \mapsto \mathbf{q}(t) = \mathbf{P}_h \mathbf{H}_L^{-1} \mathbf{A}^T \mathbf{b}(t)$. From this curve in configuration space, we determine the corresponding tangent vector field by explicitly computing the time derivative as

$$t \mapsto \dot{\mathbf{q}}(t) = \frac{d}{dt} \mathbf{q}(t). \quad (36)$$

The corresponding curve in Hamiltonian phase space is then given by

$$\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix}(t) = \begin{pmatrix} \mathbf{q}(t) \\ \mathbf{p}(\mathbf{q}(t), \dot{\mathbf{q}}(t), t) \end{pmatrix}. \quad (37)$$

The coordinate transformation $\mathbf{p}(\mathbf{q}, \dot{\mathbf{q}}, t)$ is again given by Eq. 23. Corresponding to the caveat described above, however, it is crucial that it is computed with the correct function $\dot{\mathbf{q}}(t)$ determined by Eq. 36. For evaluation of the Euler-Lagrange operator ψ , we also compute $t \mapsto \dot{\mathbf{p}}(t) = (d/dt) \mathbf{p}(t)$, and then perform functional gradient descent by using the known $t \mapsto (\mathbf{q}, \mathbf{p}, \dot{\mathbf{p}})(t)$ to obtain $t \mapsto \psi(t)$ via Eq. 33.

⁴A clear exception is when the Lagrangian L does not depend on $\dot{\mathbf{q}}$, e.g., [23].

8.3 Functional Gradient Descent to Optimal Curve

If the curve $t \mapsto (\mathbf{q}, \mathbf{p}, t)(t)$ is non-optimal, i.e., when $t \mapsto \psi(t) \neq 0$, we can converge to an optimal curve using gradient descent in function space. We refer to Appendix J, or Daßler and Günther [6], for the necessary variational background. For a given curve $t \mapsto \mathbf{q}(t)$, we now consider the curve $t \mapsto (\mathbf{q} + \boldsymbol{\eta})(t)$ with a specific *variation*

$$\boldsymbol{\eta}(t) = \frac{\delta S}{\delta \mathbf{q}}(t) = \psi(t). \quad (38)$$

Because the Euler-Lagrange operator ψ gives the functional derivative of the action $S[\mathbf{q}]$, i.e., $\delta S / \delta \mathbf{q} = \psi$, with the corresponding functional differential given by $\delta S[\mathbf{q}, \boldsymbol{\eta}]$, the operator ψ represents an infinite-dimensional gradient in function space in whose direction $\epsilon \psi$, $\epsilon > 0$, the action integral $S[\mathbf{q}]$ changes the fastest. As shown by Daßler and Günther [6], we can therefore perform gradient descent in function space, where the direction of descent is directly given by $-\psi(t)$.

We solve a boundary value problem with a *free boundary*. See in particular Appendix K.2. (One can also choose a *fixed* initial boundary.)

We now enumerate successive iterations of curves with index i . In each iteration, from the previous curve $t \mapsto \mathbf{q}_i(t)$, we compute the updated curve $t \mapsto \mathbf{q}_{i+1}(t)$ using the gradient $t \mapsto \psi_i(t)$. That is, we perform *gradient descent* with a small step size ϵ by the update rule

$$\mathbf{q}_{i+1}(t) = \mathbf{q}_i(t) - \epsilon \psi_i(t). \quad (39)$$

In order to compute the gradient $t \mapsto \psi_{i+1}(t)$ for the next iteration, all functions that $\psi(t)$ depends on need to be updated accordingly: The function $\dot{\mathbf{q}}_{i+1}(t)$ results from taking the time derivative on both sides of Eq. 39, giving $\dot{\mathbf{q}}_{i+1}(t) = \dot{\mathbf{q}}_i(t) - \epsilon (d/dt) \psi_i(t)$. Inserting into Eq. 23 then gives $\mathbf{p}_{i+1}(t)$. Combining these steps, inserting both coordinate transformations from Eqs. 23, 24, gives the direct update

$$\mathbf{p}_{i+1}(t) = \mathbf{p}_i(t) - \epsilon \left(L_{\dot{\mathbf{q}}\mathbf{q}} \psi_i + L_{\dot{\mathbf{q}}\dot{\mathbf{q}}} \frac{d}{dt} \psi_i \right)(t). \quad (40)$$

From the updated $\mathbf{p}(t)$ we then compute $\dot{\mathbf{p}}(t) = (d/dt) \mathbf{p}(t)$, and from Eq. 33 we obtain the next gradient $t \mapsto \psi_{i+1}(t)$ using $(\mathbf{q}, \mathbf{p}, \dot{\mathbf{p}})_{i+1}(t)$.

9 INTRINSICALLY STEADY VELOCITY FIELDS

We define a velocity vector field $\mathbf{v}(x, t)$ to be *intrinsically steady*, if there *exists* a curve $t \mapsto \mathbf{q}(t)$ determining a Euclidean reference frame as a Killing field, such that the Euler-Lagrange equation is fulfilled for the observed time derivative Lagrangian in Eq. 9, 10a, and we have

$$L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t) = 0, \quad \dot{\mathbf{q}} = \frac{d\mathbf{q}}{dt}. \quad \text{Thus, } S[\mathbf{q}] = 0. \quad (41)$$

In our framework, this condition arises naturally, but it is not ideally practical: One has to actually compute curves $t \mapsto \mathbf{q}(t)$ and check whether Eq. 41 is fulfilled. We therefore want to obtain a general condition that is both necessary and sufficient for a given vector field $\mathbf{v}$ to be intrinsically steady, without actually computing an optimal observer.⁵

⁵In fact, we will obtain (Theorem 1) that *if* the condition in Eq. 46 is fulfilled, the corresponding points $\mathbf{q}(t)$ used to define $t \mapsto \mathbf{q}(t)$ point-wise, will *automatically fulfill* the Euler-Lagrange equation. But we do not have to *solve* it.

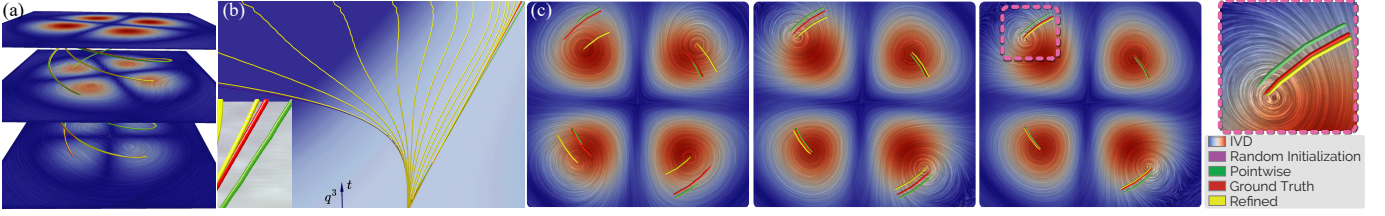


Fig. 7: **Functional gradient descent** (RFC-U flow, $k = 3$). (a) The input field $\mathbf{v}$ and its objective vortex core lines. (b) The current observer curve $t \mapsto \mathbf{q}^3(t)$ in the $\mathbf{p} = 0$ subspace. (c) Intermediate observed fields $\mathbf{v} - \mathbf{u}$ with LIC and extracted core lines during optimization. From left to right, the solution converges from an incorrect state to the correct one. Our result is more accurate than the point-wise observer curve (green).

Because zero is the smallest value of any function L given by Eq. 11, if the above curve $t \mapsto \mathbf{q}(t)$ exists, at every parameter t we must have

$$L_{\mathbf{q}} = \frac{\partial L}{\partial \mathbf{q}} = 0, \quad \text{and} \quad L_{\dot{\mathbf{q}}} = \frac{\partial L}{\partial \dot{\mathbf{q}}} = \mathbf{p} = 0. \quad (42)$$

From the general definition of the Hamiltonian H for any Lagrangian L (Eq. 17), with $\mathbf{p} = 0$ we obtain $H(\mathbf{q}, 0, t) = -L(\mathbf{q}, \dot{\mathbf{q}}(\mathbf{q}, 0, t), t)$. Thus, for any optimal curve $t \mapsto \mathbf{q}(t)$ fulfilling Eq. 41, we must have $H(\mathbf{q}(t), 0, t) = 0$ for all t . This gives the condition that the differential dH along the curve must vanish, which we can write as $dH(\bar{X}_H) = 0$, with the Hamiltonian vector field $\bar{X}_H$ corresponding to the curve $t \mapsto (\mathbf{q}(t), 0, t)(t)$. However, for *any* such field $\bar{X}_H$ (Eq. 84)

$$dH(\bar{X}_H) = \frac{\partial H}{\partial t}. \quad (43)$$

We thus obtain the condition that for an intrinsically steady input field and the corresponding optimal curve $t \mapsto \mathbf{q}(t)$, we must have

$$H(\mathbf{q}(t), 0, t) = \frac{\partial H}{\partial t}(\mathbf{q}(t), 0, t) = 0. \quad (44)$$

So far we have considered the case of an optimal curve $t \mapsto \mathbf{q}(t)$ that is known to *exist*, but we do not know it explicitly. An absolutely crucial, and as above for Eq. 35 subtle point, however, is this one: Any valid curve $t \mapsto \mathbf{q}(t)$, with $\mathbf{p}(t) = 0$, *must*, at every parameter t , fulfill

$$\dot{\mathbf{q}}(\mathbf{q}(t), 0, t) = \frac{d}{dt}\mathbf{q}(t). \quad (45)$$

If a curve $t \mapsto (\mathbf{q}(t), 0, t)(t)$ is not computed as a Hamiltonian trajectory, but, e.g., point-wise via Eq. 34, this condition is, however, in general not fulfilled: Determining $\dot{\mathbf{q}}(t)$ independently of $\mathbf{q}(t)$, which Eq. 34 does by using $2k$ independent variables $(\mathbf{q}, \dot{\mathbf{q}})$ instead of the k independent variables $\mathbf{q}$ of a curve $t \mapsto \mathbf{q}(t)$ in configuration space, will instead usually result in Eq. 35. Nevertheless, for the important *special case* of vector fields that are intrinsically steady, we can obtain

Theorem 1. *If a vector field $\mathbf{v}(x, t)$ is intrinsically steady, and the optimal curve $t \mapsto \mathbf{q}(t)$ is unique, it can in fact be computed point-wise using Eq. 34. In this specific case, Eq. 45 will always be fulfilled.*

For this theorem we assume the Lagrangian from Eq. 10a integrated over a region $U(t)$ as in Eq. 9. Uniqueness of $t \mapsto \mathbf{q}(t)$ is given when the Hessian matrix $\mathbf{H}_L$ is invertible, which is a very mild condition that is fulfilled for all vector fields that are not so simple such that infinitely many observers see them as steady (this is true for Killing fields).

Furthermore, the more fundamental observation is

Theorem 2. *With the observed time derivative Lagrangian in Eq. 9, 10a for a Euclidean reference frame modeled by a Killing vector field, if the corresponding Hessian matrix $\mathbf{H}_L(t)$ of a vector field $\mathbf{v}(x, t)$ is invertible, the field $\mathbf{v}$ is intrinsically steady if and only if, for every t ,*

$$H(\mathbf{q}, 0, t) = \frac{\partial H}{\partial t}(\mathbf{q}, 0, t) = 0, \quad \text{where } \mathbf{q} = \mathbf{P}_h \mathbf{H}_L^{-1} \mathbf{A}^T \mathbf{b}(t). \quad (46)$$

Proof. We prove Theorem 1 and Theorem 2 by proving the more general Theorem 3 in Appendix H that subsumes both as special cases. $\square$

Moreover, to be able to evaluate Eq. 46, in our framework we again do not even need to know the Hamiltonian H . See Appendix H.1.

We observe the following analogy: For a steady input field $\mathbf{v}(x, t)$, i.e., $\partial \mathbf{v} / \partial t = 0$, we have $\partial H / \partial t = 0$ everywhere. A standard result in Hamiltonian mechanics is that when H is time-independent, $dH(\bar{X}_H) = 0$ (Eq. 83). Now, for an *intrinsically steady input field* we have obtained the analogous more general condition given by Eq. 46.

10 RESULTS AND DISCUSSION

Fig. 4 demonstrates that approaches that are not integrable in time (here, [10]) do, in general, not obtain Euler-Lagrange solutions, and thus the computed observers do not have minimum cost over time. In contrast, our framework guarantees provably optimal observers.

Hamiltonian vector fields. Fig. 5 depicts several example time steps of Hamiltonian vector fields for our data sets, and the corresponding optimal observer positions in phase space. Fig. 6 visualizes the Hamiltonian vector field of the RFC-A flow. For $k = 3$, Eq. 26 defines a six-dimensional unsteady vector field in $(\mathbf{q}, \mathbf{p})$. We focus on its (q^3, p_3, t) subspace, with time shown vertically. In Fig. 6(a), we render two time slices of this subfield as stream lines, colored by the Hamiltonian H from Eq. 32. Together, these views show how the saddle structure evolves over time. We also overlay the optimal observer trajectory, providing a geometric interpretation of the solution within the Hamiltonian vector field. Fig. 6(b) restricts the view to the $\mathbf{p} = 0$ plane, where the optimal solution appears as a time-varying trajectory in q^3 . For this flow, the optimal observer has linearly accelerating angular velocity $\omega(t) = at + \omega(0)$, with $a = -0.1$ and initial value $\omega(0) = -1$. Because RFC-A is intrinsically steady, the point-wise solution and the optimal Hamiltonian trajectory coincide exactly. This agreement confirms that, for steady flows, the state variable $\mathbf{q}(t)$ directly matches the observer recovered from the point-wise formulation. In Fig. 6(c), we show additional Hamiltonian trajectories initialized at the same $q^3 = -1$ but with non-zero momenta. The optimal observer trajectory is the unique one in this family that remains on the $\mathbf{p} = 0$ plane. This view clarifies the role of the Hamiltonian formulation: The desired observer is not an isolated curve constructed point-wise, but one distinguished trajectory of the full Hamiltonian system.

Observer refinement. In Fig. 7, functional gradient descent is used to refine an observer for the RFC-U flow. We start from a randomly

Table 2: **Performance comparison.** Because our method is progressive in time, we report run time [seconds] per time step, except for Hadwiger et al. [15] over all time steps. *Ours (Killing)* is a Killing observer (2D: $k = 3$; 3D: $k = 6$); *Ours (non-Killing)* is an approximate Killing field ($k = N \cdot n$). “—” were not measured due to prohibitive memory usage.

Fig.	Data Set	Resolution $X \times Y (\times Z) \times T$	Günther [10]		Hadwiger [15]		VDE [14]		Ours	
			time step	total	time step	total	time step	total	Killing	non-Killing
6	RFC-A	$(64 \times 64) \times 64$	0.23	22.97	0.03	0.004	0.03	0.004	0.004	0.08
7	RFC-U	$(128 \times 128) \times 128$	2.17	35.50	0.13	0.004	0.13	0.004	0.004	0.08
8	Cylinder2D	$(640 \times 80) \times 100$	5.86	430	0.77	0.006	0.77	0.006	0.006	0.45
9	Cylinder3D	$(320 \times 120 \times 40) \times 151$	~45.2	—	~12.1	0.950	~12.1	0.950	0.950	8.00
11	Boussinesq	$(150 \times 450) \times 150$	12.1	870	1.27	0.008	1.27	0.008	0.008	0.86
12	pipeCylinder	$(450 \times 150) \times 1501$	8.9	—	1.39	0.005	1.39	0.005	0.005	0.95

Table 3: **Quantitative comparison** of observer computation for two data sets with known ground truth (Appendix A). We report the PSNR against the ground truth for the observer fields of different methods.

Data Set	Günther [10]	Hadwiger [15]	Point-wise	Ours (raw)	Ours (refined)
RFC-A	41.6	44.5	54.2	51.5	58.7
RFC-U	32.1	40.9	54.0	19.2	54.2

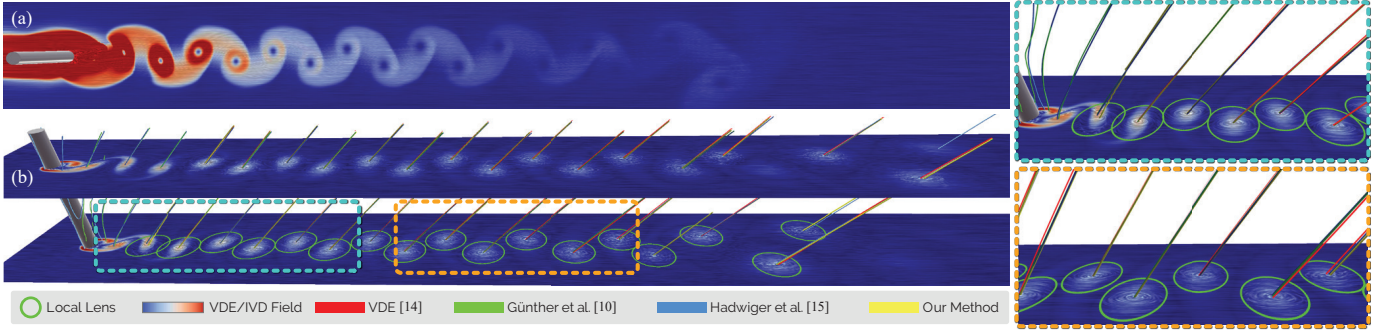


Fig. 8: **Computing different approaches in our framework (Cylinder2D).** (a) The VDE [14] scalar field computed by our framework. (b) Core lines obtained with the definitions of Günther et al. [10], Hadwiger et al. [15], and Günther and Theisel [14], in our framework.

chosen non-optimal curve (purple) in the $\mathbf{p} = 0$ subspace of the Hamiltonian phase space and iteratively update it using the functional gradient. Each intermediate solution curve (yellow) defines one observer field $\mathbf{u}$. As shown in Fig. 7(b), the trajectory is progressively refined, approaching an optimal observer until converging to an exact Euler-Lagrange solution (red). This evolution of the observer is reflected directly in the observed fields $\mathbf{v} - \mathbf{u}$ in Fig. 7(c): the LIC patterns and extracted core lines evolve from an incorrect configuration toward the expected vortex structure. In contrast to the point-wise observer curve, which accumulates larger error, our optimization enforces temporal consistency over the full trajectory and yields a more accurate observer.

Progressively computing previous methods. The generic Lagrangian of Eq. 9 allows us to compute previous observer optimization approaches in our framework, however with added features such as faster (Table 2) progressive computation and refinement toward exact Euler-Lagrange solutions. Figs. 8 and 11 demonstrate that our framework reproduces several established objective core line formulations while combining it with progressive computation in time. On the classical *Cylinder2D* wake (Fig. 8), we reproduce a VDE [14] scalar field within our framework, and recover core lines corresponding to Günther et al. [10], Hadwiger et al. [15], and Günther and Theisel [14]. For Günther et al. [10] and Hadwiger et al. [15], we first compute the observer field, then extract critical curves from the resulting observed field, followed by IVD-based filtering. For VDE [14], the observer defines a VDE scalar field, and core lines are path lines starting at its minima. The extracted curves largely coincide, indicating that these formulations are reproduced consistently in our unified setting. The Boussinesq example (Fig. 11) provides a complementary validation on a more complex flow. We verify correctness by checking whether nearby path lines rotate around the extracted core line. The LIC visualization in Fig. 11 shows the observed vector field, and the zoomed-in view confirms that our extracted core lines remain consistent with Günther et al. [10] and VDE [14]. Using the term L_x^S , we can also compute the approach by Kaszás et al. [23], which minimizes the observed kinetic

energy (Fig. 13). Additional results can be found in Appendix A.7.

Application to 3D unsteady flow. Fig. 9 compares our method with the local reference frame method of Günther et al. [10] on the *Cylinder3D* data set at $Re = 160$. Their method is a strong local baseline for objective vortex analysis, estimating a locally optimal observer at each sample point from a $(13 \times 13 \times 13)$ spatial neighborhood. In Fig. 9 (a), the upper-right inset shows an IVD volume rendering that provides context for the vortical cylinder wake. In contrast, our method computes locally optimal Killing reference frames using the proposed Hamiltonian formulation with functional gradient descent. The inset in Fig. 9 (b) illustrates the selected local region. In the yellow highlighted wake region, our method extracts more stable and coherent local vortical structures, while the baseline result is more fragmented.

Performance evaluation. Table 2 compares the runtime of our progressive framework with prior baselines. Since our method is progressive in time, all columns report per-time step optimization time, except for Hadwiger et al. [15], which reports total run time over the full spacetime domain. Our approximate Killing fields have the same number of spatial unknowns as prior work [15], but we have removed the need for huge matrices containing all time steps. Our Killing fields have a very small problem size. These examples show that our framework matches prior methods qualitatively while achieving substantially lower run times. Performance was measured on a workstation with two Intel(R) Xeon(R) Gold 6230R CPUs @ 2.10GHz. For Günther et al. [10], we use local size = 13; for VDE [14], we use $\tau = 1.0$ and $R = 11$. For Killing observers, $\lambda_K = \lambda_S = 0$. For approximate Killing observers, we use $\lambda_K = 0.5$ and $\lambda_S = 0.1$ for all experiments. Our code is publicly available at <https://github.com/Cindy-xdZhang/GenericVariationalVortexCore>.

Quantitative evaluation. Table 3 compares the PSNR of different observer computation strategies relative to known ground truth. Point-wise computation is by Eq. 34 at each time step. Our raw result is a Killing field that is then further refined by functional gradient descent.

11 CONCLUSIONS

A strength of our variational framework for practical applications is that it not only unifies previous approaches for observer optimization but also allows adding novel algorithmic features to them. In particular, it enables progressive computation in time for all methods, for faster computation but also for applications such as vortex lens self-advection in time [32]. Non-optimal observers computed via any method can be refined toward an exact solution of the Euler-Lagrange equation. We provide a strong foundation for future work, by enabling rigorous comparison and analysis of different approaches in a unified mathematical setting. Additional Lagrangians can easily be defined and analyzed.

On the other hand, an important theoretical contribution of our framework is that *given a specific choice of Lagrangian L* , the corresponding Hamiltonian vector field is fully determined, and therefore *all possible optimal observers are determined*. A crucial goal in practice is therefore *choosing a particular optimal observer* via a suitable strategy.

Our definition of intrinsically steady vector fields also shows that our Hamiltonian phase space furthermore reflects intrinsic properties of flow fields themselves, not just the properties of specific observers.

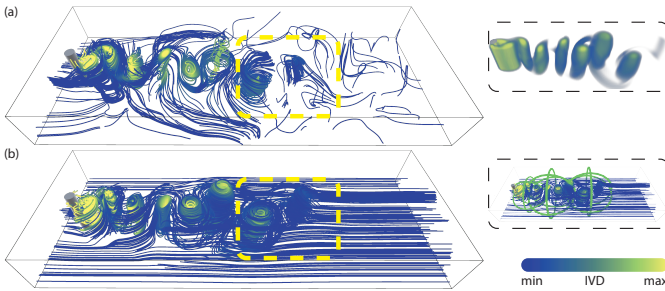


Fig. 9: **Cylinder3D ($Re = 160$) results.** (a) Local observer optimization of Günther et al. [10] vs. (b) our result computed in the selected region. Stream lines are color-coded by IVD. Insets show the IVD volume context in (a) and the selected local region in (b). Both results use the exact same stream line seeding and integration settings; therefore, the visual differences reflect the different observer fields. Our method extracts more stable vortical structures in the highlighted wake region.

ACKNOWLEDGMENTS

This work was supported by King Abdullah University of Science and Technology (KAUST) baseline funding.

REFERENCES

- [1] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer, 2nd ed., 1989. 1, 3, 5, 13
- [2] G. Astarita. Objective and generally applicable criteria for flow classification. *Journal of Non-Newtonian Fluid Mechanics*, 6(1):69–76, 1979. doi: [10.1016/0377-0257\(79\)87004-4](https://doi.org/10.1016/0377-0257(79)87004-4) 2
- [3] I. Baeza Rojo and T. Günther. Vector field topology of time-dependent flows in a steady reference frame. *IEEE Transactions on Visualization and Computer Graphics (Proceedings IEEE VIS 2019)*, 26(1):280–290, 2020. doi: [10.1109/TVCG.2019.2934375](https://doi.org/10.1109/TVCG.2019.2934375) 2
- [4] R. Bujack, M. Hlawitschka, and K. I. Joy. Topology-inspired Galilean invariant vector field analysis. In *Proceedings of IEEE Pacific Visualization 2016*, pp. 72–79, 2016. doi: [10.1109/PACIFICVIS.2016.7465253](https://doi.org/10.1109/PACIFICVIS.2016.7465253) 2
- [5] R. Bujack, L. Yan, I. Hotz, C. Garth, and B. Wang. State of the art in time-dependent flow topology: Interpreting physical meaningfulness through mathematical properties. *Computer Graphics Forum*, 39(3):811–835, 2020. doi: [10.1111/cgf.14037](https://doi.org/10.1111/cgf.14037) 2
- [6] N. Dassler and T. Günther. Variational feature extraction in scientific visualization. *ACM Transactions on Graphics*, 43(4):1–16, July 2024. doi: [10.1145/3658219](https://doi.org/10.1145/3658219) 7
- [7] R. Drouot and M. Lucius. Approximation du second ordre de la loi de comportement des fluides simples. lois classiques déduites de l'introduction d'un nouveau tenseur objectif. *Archiwum Mechaniki Stosowanej*, 28(2):189–198, 1976. 2
- [8] I. M. Gelfand and S. V. Fomin. *Calculus of Variations*. Prentice-Hall, 1963. 1, 3, 4, 12
- [9] H. Goldstein, C. P. Poole, and J. L. Safko. *Classical Mechanics*. Addison-Wesley, 3rd ed., 2001. 1, 3, 4, 5, 6, 12
- [10] T. Günther, M. Gross, and H. Theisel. Generic objective vortices for flow visualization. *ACM Transactions on Graphics*, 36(4), July 2017. doi: [10.1145/3072959.3073684](https://doi.org/10.1145/3072959.3073684) 1, 2, 4, 5, 8, 9, 11, 12
- [11] T. Günther, M. Schulze, and H. Theisel. Rotation invariant vortices for flow visualization. *IEEE Transactions on Visualization and Computer Graphics (Proceedings IEEE VIS 2015)*, 22(1):817–826, 2016. doi: [10.1109/TVCG.2015.2467200](https://doi.org/10.1109/TVCG.2015.2467200) 2
- [12] T. Günther and H. Theisel. The state of the art in vortex extraction. *Computer Graphics Forum*, 37(6):149–173, 2018. doi: [10.1111/cgf.13319](https://doi.org/10.1111/cgf.13319) 2
- [13] T. Günther and H. Theisel. Hyper-objective vortices. *IEEE Transactions on Visualization and Computer Graphics*, 26(3):1532 – 1547, 2020. doi: [10.1109/TVCG.2018.2868760](https://doi.org/10.1109/TVCG.2018.2868760) 2
- [14] T. Günther and H. Theisel. Objective lagrangian vortex cores and their visual representations. *IEEE Transactions on Visualization and Computer Graphics (Proceedings IEEE VIS 2024)*, 31(1):76–85, 2025. doi: [10.1109/TVCG.2024.3456384](https://doi.org/10.1109/TVCG.2024.3456384) 1, 2, 3, 4, 8, 9, 11, 12, 20
- [15] M. Hadwiger, M. Mlejnek, T. Theußl, and P. Rautek. Time-dependent flow seen through approximate observer killing fields. *IEEE Transactions on Visualization and Computer Graphics (Proceedings IEEE VIS 2018)*, 25(1):1257–1266, 2019. doi: [10.1109/TVCG.2018.2864839](https://doi.org/10.1109/TVCG.2018.2864839) 1, 2, 3, 4, 8, 9, 12
- [16] M. Hadwiger, T. Theußl, and P. Rautek. Riemannian geometry for scientific visualization. *SIGGRAPH Asia 2022 Courses*, (5), 2022. doi: [10.1145/3550495.3558227](https://doi.org/10.1145/3550495.3558227) 4, 13
- [17] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer, 2nd ed., 2006. 15
- [18] G. Haller. An objective definition of a vortex. *Journal of Fluid Mechanics*, 525:1–26, 2005. doi: [10.1017/S0022112004002526](https://doi.org/10.1017/S0022112004002526) 2
- [19] G. Haller. Lagrangian coherent structures. *Annual review of fluid mechanics*, 47:137–162, 2015. doi: [10.1146/annurev-fluid-010313-141322](https://doi.org/10.1146/annurev-fluid-010313-141322) 2
- [20] G. Haller, A. Hadjighasem, M. Farazmand, and F. Huhn. Defining coherent vortices objectively from the vorticity. *Journal of Fluid Mechanics*, 795:136–173, 2016. doi: [10.1017/jfm.2016.151](https://doi.org/10.1017/jfm.2016.151) 2
- [21] G. Haller and G. Yuan. Lagrangian coherent structures and mixing in two-dimensional turbulence. *Physica D: Nonlinear Phenomena*, 147(3):352–370, 2000. doi: [10.1016/S0167-2789\(00\)00142-1](https://doi.org/10.1016/S0167-2789(00)00142-1) 2
- [22] J. Kasten, J. Reininghaus, I. Hotz, and H.-C. Hege. Two-dimensional time-dependent vortex regions based on the acceleration magnitude. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2080–2087, 2011. doi: [10.1109/TVCG.2011.249](https://doi.org/10.1109/TVCG.2011.249) 2
- [23] B. Kaszás, T. Pedernana, and G. Haller. The objective deformation component of a velocity field. *European Journal of Mechanics - B/Fluids*, 2022. doi: [10.1016/j.euromechflu.2022.12.007](https://doi.org/10.1016/j.euromechflu.2022.12.007) 2, 4, 7, 9, 11
- [24] B. Kim and T. Günther. Robust reference frame extraction from unsteady 2d vector fields with convolutional neural networks. *Computer Graphics Forum*, 38(3):285–295, 2019. doi: [10.1111/cgf.13689](https://doi.org/10.1111/cgf.13689) 2
- [25] C. Lanczos. *The Variational Principles of Mechanics*. Dover Publications, Inc., 1986. 1, 13
- [26] D. Lovelock and H. Rund. *Tensors, Differential Forms, and Variational Principles*. Dover Publications, Inc., 1975. 22, 23
- [27] H. J. Lugt. The dilemma of defining a vortex. In *Recent developments in theoretical and experimental fluid mechanics*, pp. 309–321. Springer, 1979. doi: [10.1007/978-3-642-67220-0_32](https://doi.org/10.1007/978-3-642-67220-0_32) 2
- [28] P. Mann. *Lagrangian and Hamiltonian Dynamics*. Oxford University Press, 2018. 3
- [29] R. W. Ogden. *Non-Linear Elastic Deformations*. Dover Publications, Inc., 1997. 2
- [30] F. H. Post, B. Vrolijk, H. Hauser, R. S. Laramée, and H. Doleisch. The state of the art in flow visualisation: Feature extraction and tracking. *Computer Graphics Forum*, 22(4):775–792, 2003. doi: [10.1111/j.1467-8659.2003.00723.x](https://doi.org/10.1111/j.1467-8659.2003.00723.x) 2
- [31] P. Rautek, M. Mlejnek, J. Beyer, J. Troidl, H. Pfister, T. Theußl, and M. Hadwiger. Objective observer-relative flow visualization in curved spaces for unsteady 2D geophysical flows. *IEEE Transactions on Visualization and Computer Graphics (Proceedings IEEE VIS 2020)*, 27(2):283–293, 2021. doi: [10.1109/TVCG.2020.3030454](https://doi.org/10.1109/TVCG.2020.3030454) 1, 2, 3, 4, 12
- [32] P. Rautek, X. Zhang, B. Woschizka, T. Theußl, and M. Hadwiger. Vortex lens: Interactive vortex core line extraction using observed line integral convolution. *IEEE Transactions on Visualization and Computer Graphics (Proceedings IEEE VIS 2023)*, 30(1):55–65, 2024. doi: [10.1109/TVCG.2023.3326915](https://doi.org/10.1109/TVCG.2023.3326915) 1, 2, 4, 9
- [33] J. Rey Ramirez, P. Rautek, T. Günther, and M. Hadwiger. Locally adapted reference frame fields using moving least squares. *IEEE Transactions on Visualization and Computer Graphics (Proceedings IEEE VIS 2025)*, 32(1):1–10, 2026. doi: [10.1109/TVCG.2025.3634845](https://doi.org/10.1109/TVCG.2025.3634845) 1, 2, 4, 12
- [34] J. Sahner, T. Weinkauff, and H.-C. Hege. Galilean invariant extraction and iconic representation of vortex core lines. In *Proceedings of Eurographics/IEEE VGTC Symposium on Visualization*, pp. 151–160, 2005. 2
- [35] S. C. Shadden, F. Lekien, and J. E. Marsden. Definition and properties of lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows. *Physica D: Nonlinear Phenomena*, 212(3):271–304, 2005. doi: [10.1016/j.physd.2005.10.007](https://doi.org/10.1016/j.physd.2005.10.007) 2
- [36] H. Theisel, M. Hadwiger, P. Rautek, T. Theußl, and T. Günther. Vortex criteria can be objectivized by unsteadiness minimization. *Physics of Fluids*, 33(10):107115, 2021. doi: [10.1063/5.0063817](https://doi.org/10.1063/5.0063817) 2
- [37] H. Theisel and H.-P. Seidel. Feature flow fields. In G.-P. Bonneau, S. Hahmann, and C. H. (Ed.), eds., *Proc. Joint Eurographics - IEEE TCVG Symposium on Visualization (VisSym'03)*, pp. 141–148, 2003. 2
- [38] C. Truesdell and W. Noll. *The Nonlinear Field Theories of Mechanics*. Springer, 1965. 2
- [39] T. Weinkauff, J. Sahner, H. Theisel, and H.-C. Hege. Cores of swirling particle motion in unsteady flows. *IEEE Transactions on Visualization and Computer Graphics (Proceedings IEEE Visualization)*, 13(6):1759–1766, 2007. doi: [10.1109/TVCG.2007.70545](https://doi.org/10.1109/TVCG.2007.70545) 2
- [40] X. Zhang, A. Ageeli, T. Theußl, M. Hadwiger, and P. Rautek. Exploring 3d unsteady flow using 6d observer space interactions. *IEEE Transactions on Visualization and Computer Graphics (Proceedings IEEE VIS 2025)*, 32(1):1–10, 2026. doi: [10.1109/TVCG.2025.3642506](https://doi.org/10.1109/TVCG.2025.3642506) 2, 3, 4
- [41] X. Zhang, M. Hadwiger, T. Theußl, and P. Rautek. Interactive exploration of physically-observable objective vortices in unsteady 2D flow. *IEEE Transactions on Visualization and Computer Graphics (Proceedings IEEE VIS 2021)*, 28(1):281–290, 2022. doi: [10.1109/TVCG.2021.3115565](https://doi.org/10.1109/TVCG.2021.3115565) 2, 3, 4

A DATA SETS

We describe the unsteady flow fields used in the paper.

A.1 RFC-C (Four Centers in Constantly Rotating Frame)

This unsteady vector field has originally been defined by Günther et al. [10]. Using Cartesian coordinates (x, y) in the spatial domain $U = [-2, 2] \times [-2, 2]$, it is given in the time range $[t_a, t_b] = [0, 2\pi]$ by

$$\mathbf{v}(x, y, t) = \begin{pmatrix} \cos \omega t & \sin \omega t \\ -\sin \omega t & \cos \omega t \end{pmatrix} \begin{pmatrix} \bar{u}(t) + \omega v(t) \\ \bar{v}(t) - \omega u(t) \end{pmatrix}. \quad (47)$$

For brevity, above the dependence of the functions $\bar{u}(t)$, $\bar{v}(t)$, $u(t)$, $v(t)$ on the coordinates x, y is implicit, and they are defined by

$$\begin{aligned} \bar{u}(t) &= -u(t)(2v(t)^2 - 1)e^{-u(t)^2 - v(t)^2}, \\ \bar{v}(t) &= v(t)(2u(t)^2 - 1)e^{-u(t)^2 - v(t)^2}, \end{aligned} \quad (48)$$

with

$$\begin{aligned} u(t) &= x \cos \omega t - y \sin \omega t, \\ v(t) &= x \sin \omega t + y \cos \omega t. \end{aligned} \quad (49)$$

This vector field is derived from an originally steady input field which is viewed in a reference frame that is rotating with constant angular velocity $\omega = 1$. It is therefore an *intrinsically steady* field.

The optimal observer field $\mathbf{u}$ minimizing the observed time derivative $L_{\mathbf{u}}(\mathbf{v} - \mathbf{u})$ can be given analytically. In Cartesian coordinates, it is

$$\mathbf{u}(x, y, t) = -\omega \begin{pmatrix} -y \\ x \end{pmatrix}. \quad (50)$$

We can analytically verify that with this $\mathbf{u}$, indeed, $L_{\mathbf{u}}(\mathbf{v} - \mathbf{u}) = \partial \mathbf{v} / \partial t - \partial \mathbf{u} / \partial t + \nabla \mathbf{v}(\mathbf{u}) - \nabla \mathbf{u}(\mathbf{v}) = 0$, for all (x, y, t) . In our framework, the optimal observer is therefore given by ($\omega = 1$)

$$\begin{aligned} \mathbf{q}(t) &= (0, 0, -1)^T, \\ \dot{\mathbf{q}}(t) &= (0, 0, 0)^T. \end{aligned} \quad (51)$$

A.2 RFC-A (Four Centers in Rotation-Accelerating Frame)

We define a version of the RFC-C data set where, instead of with constant angular velocity, the frame is rotating with linearly accelerating angular velocity $\omega(t) = at + \omega(0)$, for a constant $a = -0.1$, and $\omega(0) = -1$.

In our framework, the optimal observer is therefore given by

$$\begin{aligned} \mathbf{q}(t) &= (0, 0, -\omega(t))^T, \\ \dot{\mathbf{q}}(t) &= (0, 0, -a)^T. \end{aligned} \quad (52)$$

As the RFC-C flow, the RFC-A flow is also an *intrinsically steady* field.

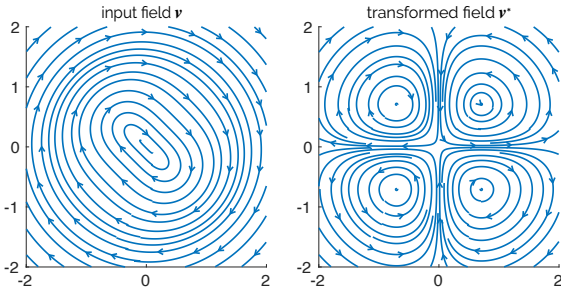


Fig. 10: (Left) Unsteady RFC-C flow field, given by Eq. 47 ($t = 0$). (Right) Steady observed velocity field in the optimal reference frame.

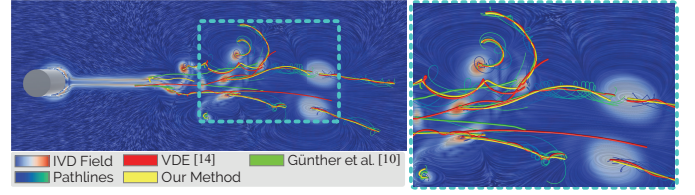


Fig. 11: Computing different approaches in our framework (*Boussinesq*). As in Fig. 8, different Lagrangians in our framework replicate the results of Günther et al. [10], and Günther and Theisel [14] (VDE).

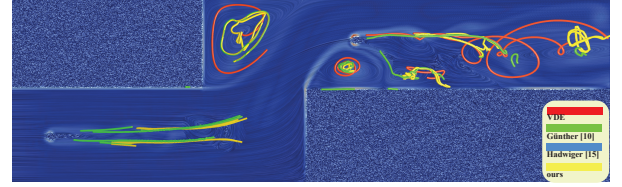


Fig. 12: Computing core lines using different approaches in our framework (*pipeCylinder* data set).

A.3 RFC-U (Four Centers, Intrinsically Unsteady)

We define an *intrinsically unsteady* field $\tilde{\mathbf{v}}$ from the field in Eq. 47, however with angular velocity $\omega(t)$ as in the RFC-A flow, by defining

$$\tilde{\mathbf{v}}(x, y, t) = \begin{pmatrix} \cos \omega t & \sin \omega t \\ -\sin \omega t & \cos \omega t \end{pmatrix} \begin{pmatrix} s(t) \cdot \bar{u}(t) + \omega v(t) \\ s(t) \cdot \bar{v}(t) - \omega u(t) \end{pmatrix}. \quad (53)$$

The vector field $\tilde{\mathbf{v}}$ will be *intrinsically unsteady* for any time-dependent scale function $s(t) \in \mathbb{R}$ for which $(d/dt)s(t) = \dot{s}(t) \neq 0$. The function $s(t)$ regulates the intrinsic vorticity of the four vortices in the RFC-C flow. However, the intrinsic unsteadiness of this field is structurally very simple, and the optimal observer is again given by

$$\begin{aligned} \mathbf{q}(t) &= (0, 0, -\omega(t))^T, \\ \dot{\mathbf{q}}(t) &= (0, 0, -\dot{\omega})^T. \end{aligned} \quad (54)$$

A.4 Cylinder2D

This is a 2D Navier-Stokes simulation of the von Kármán vortex street developing behind a round cylindrical obstacle.

A.5 Cylinder3D

This is a 3D Navier-Stokes simulation of the von Kármán vortex street developing behind a round cylindrical obstacle for $Re = 160$.

A.6 Boussinesq

This is a Navier-Stokes simulation of the flow around a heated cylinder with a Boussinesq approximation. See Fig. 11.

A.7 PipeCylinder

This is a Navier-Stokes simulation of the flow passing a sharp corner inside a pipe. See Fig. 12.

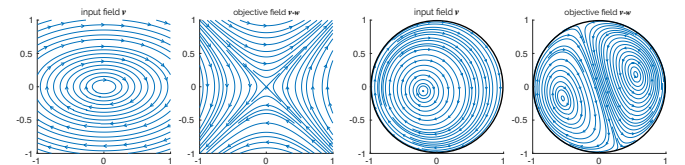


Fig. 13: (Left two images) Velocity field from [23, Eq.33] at time $t = 0$, with $\mathbf{u}$ computed using our Lagrangian L_x^S . (Right two images) Model velocity field from [23, Eq.35] at time $t = \pi/10$, with $\mathbf{u}$ computed using our Lagrangian L_x^S . For both fields, we obtain identical results to [23], with $\mathbf{u} = \mathbf{v}_{RB}$, and $\mathbf{v} - \mathbf{u}$ giving the *objective deformation component* of $\mathbf{v}$, as defined by [23].

B INTEGRABILITY OF OBSERVER FIELDS

Following Rey Ramirez et al. [33], their approach and the approach of Günther et al. [10] can be seen to compute a reference frame field for unsteady flow as a field of parameters given by a function

$$\begin{aligned} \mathbf{f}: U &\rightarrow \mathbb{R}^{2k}, \\ x &\mapsto \mathbf{f}(x). \end{aligned} \quad (55)$$

For example, in 2D the frame parameters $\mathbf{f}(x)$ at a point $x \in U$ are given by $2k = 6$ independent parameters, i.e., degrees of freedom,

$$\mathbf{f}(x) = (q^1 \ q^2 \ q^3 \ \dot{q}^1 \ \dot{q}^2 \ \dot{q}^3)^T(x). \quad (56)$$

The parameters $\mathbf{f}(x)$ for a specific position x specify an entire Killing field $\mathbf{u}$ as well as its time derivative over the domain U , as given by Eq. 2 and Eq. 6, respectively. However, this means that these approaches in fact use $2k$ degrees of freedom per spatial point x , where an observer field in n dimensions has at most n degrees of freedom per point x . In $\mathbb{R}^n$, a Killing field is fully determined by k parameters ($k = 3$ for $n = 2$, $k = 6$ for $n = 3$), which is equivalent to a Killing field $\mathbf{u}(y, t)$ being fully determined by $\mathbf{u}(x, t)$ and $\nabla \mathbf{u}(x, t)$, at a single fixed point x . The entire Killing field can therefore be computed by integrating the directional derivative $\nabla \mathbf{u}$, which in $\mathbb{R}^n$ gives

$$\mathbf{u}(y, t) = \nabla \mathbf{u}(x, t) \cdot (y - x) + \mathbf{u}(x, t). \quad (57)$$

The above equation is the integral of the directional derivative $\nabla \mathbf{u}$ from the point x to the point y , with constant of integration $\mathbf{u}(x, t)$. In $\mathbb{R}^n$, the integral is immediately given in closed form as above. (On curved manifolds, e.g., Killing fields on the sphere $\mathbb{S}^2$, this integral cannot be computed as in flat $\mathbb{R}^n$.) Now, however, because the parameters $q^i(x)$ are determined *independently* at every point x , in general we will have

$$\tilde{\mathbf{u}}(\tilde{x}, t) \neq \mathbf{u}(y, t). \quad (58)$$

Here, the vector $\tilde{\mathbf{u}}(\tilde{x}, t)$ at the point $\tilde{x}$ is given by the parameters $\mathbf{f}(\tilde{x})$ at the point $\tilde{x}$, and the vector $\mathbf{u}(y, t)$, with $y = \tilde{x}$, is given via Eq. 57 and the parameters $\mathbf{f}(x)$ at the point x , with $x \neq \tilde{x}$. In general, we will have Eq. 58. Therefore, because too many degrees of freedom are used in Eq. 55, the resulting observer field is, in general, not integrable in x .

A similar non-integrability property holds for time t : Because the time derivatives $\dot{q}^i$ are also determined as *independent* parameters, i.e., degrees of freedom, in Eq. 56, in general we will also have

$$\dot{q}^i(x, t) \neq \frac{d}{dt} q^i(x, t). \quad (59)$$

This not only means that the integral of the functions $\dot{q}^i(t)$ will in general not match the functions $q^i(t)$, but also that the corresponding observer field will in general not solve the Euler-Lagrange equation.

C GENERAL FORM OF QUADRATIC LAGRANGIANS

In our framework, we can exploit that our Lagrangian L in Eq. 9 is always quadratic, with the individual quadratic terms given in Eq. 10. More specifically, all our variants of L can be written in the general form given by Eq. 11, which for convenience we repeat here as

$$L(\mathbf{q}, \dot{\mathbf{q}}, t) = \frac{1}{2} \left\| \mathbf{A}(t) \begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix} - \mathbf{b}(t) \right\|^2.$$

In this general form, $\mathbf{A}(t)$ is an $m \times 2k$ matrix, giving a linear operator $\mathbf{A}: \mathbb{R}^{2k} \rightarrow \mathbb{R}^m$, and $\mathbf{b}(t)$ is an m -dimensional vector. The number k is the dimensionality of the configuration space, and m is a constant that depends on the specific definition of L . The most important point here is that the vector whose squared length is measured by the Lagrangian L is linear (precisely, *affine*) in the form given above.

To facilitate taking derivatives, we can expand Eq. 11 as

$$\begin{aligned} L(\mathbf{q}, \dot{\mathbf{q}}, t) = & \frac{1}{2} \begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix}^T \mathbf{A}^T \mathbf{A}(t) \begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix} - \begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix}^T \mathbf{A}^T \mathbf{b}(t) + \frac{1}{2} \mathbf{b}^T \mathbf{b}(t). \end{aligned} \quad (60)$$

Now, for any fixed time t , the $2k$ first-order partial derivatives⁶ of a Lagrangian L in the form of Eq. 11 therefore have the general form

$$\left(\frac{\partial L}{\partial \mathbf{q}} \right)_{(\mathbf{q}, \dot{\mathbf{q}}, t)} = \mathbf{A}^T \mathbf{A}(t) \begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix} - \mathbf{A}^T \mathbf{b}(t). \quad (61)$$

If from this we now again compute all partial derivatives with respect to $\mathbf{q}, \dot{\mathbf{q}}$, we obtain the general form of the $2k \times 2k$ Hessian matrix

$$\mathbf{H}_L(\mathbf{q}, \dot{\mathbf{q}}, t) = \begin{pmatrix} \frac{\partial^2 L}{\partial q^i \partial q^j} & \frac{\partial^2 L}{\partial q^i \partial \dot{q}^j} \\ \frac{\partial^2 L}{\partial \dot{q}^i \partial q^j} & \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \end{pmatrix}_{(\mathbf{q}, \dot{\mathbf{q}}, t)} = \mathbf{A}^T \mathbf{A}(t) =: \mathbf{H}_L(t). \quad (62)$$

Due to L being quadratic, in the general form of Eq. 11, the Hessian matrix $\mathbf{H}_L$ is therefore *independent of the generalized coordinates* $\mathbf{q}, \dot{\mathbf{q}}$, and only depends on time t , specifically on the choice of $\mathbf{A}(t)$ in Eq. 11.

For brevity in our equations, we introduce the following short-hand notation for the four different matrix sub-blocks of the Hessian $\mathbf{H}_L$:

$$\mathbf{H}_L(t) = \begin{pmatrix} \frac{\partial^2 L}{\partial q^i \partial q^j} & \frac{\partial^2 L}{\partial q^i \partial \dot{q}^j} \\ \frac{\partial^2 L}{\partial \dot{q}^i \partial q^j} & \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \end{pmatrix}_{(t)} =: \begin{pmatrix} L_{\mathbf{q}\mathbf{q}}(t) & L_{\mathbf{q}\dot{\mathbf{q}}}(t) \\ L_{\dot{\mathbf{q}}\mathbf{q}}(t) & L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}(t) \end{pmatrix}. \quad (63)$$

As for any Hessian matrix, we can also exploit that

$$L_{\dot{\mathbf{q}}\mathbf{q}}(t) = L_{\mathbf{q}\dot{\mathbf{q}}}^T(t), \quad L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}(t) = L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}^T(t), \quad L_{\mathbf{q}\dot{\mathbf{q}}}(t) = L_{\dot{\mathbf{q}}\mathbf{q}}^T(t). \quad (64)$$

D EXPANDING THE EULER-LAGRANGE OPERATOR

Expanding the Euler-Lagrange operator $\psi(t)$ in Eq. 14 more explicitly makes clear why it comprises a system of second-order ODEs. By using the property that the tangent vector field of the curve $t \mapsto (\mathbf{q}, \dot{\mathbf{q}}, t)(t)$ is given by $t \mapsto (\dot{\mathbf{q}}, \ddot{\mathbf{q}}, 1)(t)$ we can see that we can expand Eq. 14 as

$$\psi(t) = \frac{\partial L}{\partial \mathbf{q}}(\mathbf{q}, \dot{\mathbf{q}}, t) - \begin{pmatrix} L_{\dot{\mathbf{q}}\mathbf{q}}(t) & L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}(t) \end{pmatrix} \frac{\partial^2 L}{\partial t \partial \dot{\mathbf{q}}}(\mathbf{q}, \dot{\mathbf{q}}, t) \begin{pmatrix} \dot{\mathbf{q}} \\ \ddot{\mathbf{q}} \\ 1 \end{pmatrix}. \quad (65)$$

For brevity, here and in Eq. 66 below, we abbreviate $\mathbf{q} = \mathbf{q}(t)$, $\dot{\mathbf{q}} = \dot{\mathbf{q}}(t)$, and $\ddot{\mathbf{q}} = \ddot{\mathbf{q}}(t)$. However, it is crucial that in these equations $\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}$ all correspond to a point on a specific curve $t \mapsto (\mathbf{q}, \dot{\mathbf{q}}, t)(t)$, corresponding to a particular parameter value t . The second-order derivatives comprising the $k \times k$ matrix blocks $L_{\dot{\mathbf{q}}\mathbf{q}}(t)$ and $L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}(t)$ used in Eq. 65 are given in Eq. 63. Assuming that $L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}(t)$ is invertible, we can thus fulfill the Euler-Lagrange equation $\psi = 0$ at a point $(\mathbf{q}(t), \dot{\mathbf{q}}(t), t)$, for specific t , by solving for the second-order unknown

$$\ddot{\mathbf{q}} = L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}^{-1}(t) \left(\frac{\partial L}{\partial \mathbf{q}}(\mathbf{q}, \dot{\mathbf{q}}, t) - L_{\dot{\mathbf{q}}\mathbf{q}}(t) \dot{\mathbf{q}} - \frac{\partial^2 L}{\partial t \partial \dot{\mathbf{q}}}(\mathbf{q}, \dot{\mathbf{q}}, t) \right). \quad (66)$$

It is usually a basic requirement in the calculus of variations that $L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}(t)$ is invertible in order to be able to solve the Euler-Lagrange equation [8].

However, it is an important property of our Lagrangian-Hamiltonian framework that instead of solving the system of second-order ODEs given by Eq. 66, in our framework we instead consider a *Hamiltonian phase space* and solve the corresponding first-order *Hamilton's equations* [9] there. This is simpler and faster, but completely equivalent to solving the Euler-Lagrange equation in the form of Eq. 66.

E OBJECTIVITY OF OUR VARIATIONAL FRAMEWORK

Our Lagrangian in the form of Eq. 9 is a weighted sum of the terms in Eqs. 10a, 10b, and 10c. In our framework, these terms evaluate to quantities that all have been proved in prior work to be objective [14, 15, 31]. Because adding multiple objective scalars (here, squared magnitudes of objective vectors) yields an objective scalar, all our Lagrangians in the form of Eq. 9 with Eqs. 10a, 10b, and 10c are objective.

⁶We could call this collection of $2k$ partial derivatives a “gradient.” However, $\partial L / \partial \mathbf{q}$ is in fact not tensorial (see Appendix M), and so this is not entirely correct when one assumes that the term *gradient* implies the tensor property.

An additional important question is that, even if the Lagrangian L is objective, and thus frame-indifferent, could the computed curve $t \mapsto \mathbf{q}(t)$ parameterizing the optimal observer using a basis $\mathbf{e}_i$ in fact give a different observer for different choice of basis fields $\mathbf{e}_i$? Fortunately, a fundamental property of the calculus of variations guarantees that this computation introduces no coordinate or basis dependence. As discussed in the book by Lanczos [25], the solutions of the Euler-Lagrange equation are invariant under point transformations of the configuration space (i.e., under transformations of the q^i coordinates), even if these point transformations are time-dependent. For our framework, this means that the choice of the basis $\mathbf{e}_i(x, t)$ (time-dependent or not) does not change the computed optimal observers. Thus, because the Lagrangians L are objective, we will always obtain the same optimal observer, independent of the “background reference frame” used for computation, as well as independent of the choice of basis fields $\mathbf{e}_i$.

F DERIVATION OF HAMILTON’S EQUATIONS

Given a scalar Lagrangian function L , in the Lagrangian phase space we first define a new corresponding scalar field given by the energy

$$E_L(\mathbf{q}, \dot{\mathbf{q}}, t) = \sum_{i=1}^k \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i - L = \mathbf{p}^T \dot{\mathbf{q}} - L. \quad (67)$$

From this energy specified in Lagrangian phase space, the corresponding Hamiltonian H in Hamiltonian phase space is determined by

$$\begin{aligned} H(\mathbf{q}, \mathbf{p}, t) &= E_L(\mathbf{q}, \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t), t), \\ &= p_i \dot{q}^i(\mathbf{q}, \mathbf{p}, t) - L(\mathbf{q}, \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t), t), \\ &= \mathbf{p}^T \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t) - L(\mathbf{q}, \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t), t). \end{aligned} \quad (68)$$

Due to the above Legendre transformation [1], for each time $t = \tau$,

$$L(\mathbf{q}, \dot{\mathbf{q}}, \tau) + H(\mathbf{q}, \mathbf{p}, \tau) = \mathbf{p}^T \dot{\mathbf{q}}. \quad (69)$$

We can write this with explicit arguments, viewing $\mathbf{q}, \mathbf{p}$ as independent variables, i.e., focusing on the Hamiltonian phase space $(\mathbf{q}, \mathbf{p}, t)$,

$$L(\mathbf{q}, \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, \tau), \tau) + H(\mathbf{q}, \mathbf{p}, \tau) = \mathbf{p}^T \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, \tau). \quad (70)$$

Taking partial derivatives on each side, with respect to $\mathbf{q}$ (first row below), and with respect to $\mathbf{p}$ (second row below), respectively, gives

$$\frac{\partial L}{\partial \mathbf{q}} + \frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{q}} + \frac{\partial H}{\partial \mathbf{q}} = \mathbf{p} \cdot \frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{q}}, \quad \text{and} \quad (71)$$

$$\frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{p}} + \frac{\partial H}{\partial \mathbf{p}} = \dot{\mathbf{q}} + \mathbf{p} \cdot \frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{p}}. \quad (72)$$

Canceling terms gives, for each of these equations, respectively,

$$\frac{\partial H}{\partial \mathbf{q}} = -\frac{\partial L}{\partial \mathbf{q}}, \quad \text{and} \quad (73)$$

$$\frac{\partial H}{\partial \mathbf{p}} = \dot{\mathbf{q}}. \quad (74)$$

The Euler-Lagrange operator ψ can now be written as

$$\begin{aligned} \psi &= \frac{\partial L}{\partial \mathbf{q}} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{q}}}, \\ &= -\frac{\partial H}{\partial \mathbf{q}} - \frac{d\mathbf{p}}{dt} = -\frac{\partial H}{\partial \mathbf{q}} - \dot{\mathbf{p}}. \end{aligned} \quad (75)$$

Therefore, the Euler-Lagrange equation $\psi = 0$ now gives

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}} - \psi = -\frac{\partial H}{\partial \mathbf{q}}. \quad (76)$$

Eqs. 74, 76 are *Hamilton’s equations*, forming a system of $2k$ first-order ODEs. Jointly solving both is equivalent to solving the Euler-Lagrange equation, which is a system of k second-order ODEs (Eq. 66).

That is, to determine an optimal observer given by the curve $t \mapsto (\mathbf{q}, \mathbf{p}, t)(t)$, we must solve the two vector-valued first-order ODEs

$$\begin{aligned} \dot{\mathbf{q}}(t) &= \frac{d\mathbf{q}(t)}{dt} = \frac{\partial H}{\partial \mathbf{p}}(\mathbf{q}(t), \mathbf{p}(t), t), \\ \dot{\mathbf{p}}(t) &= \frac{d\mathbf{p}(t)}{dt} = -\frac{\partial H}{\partial \mathbf{q}}(\mathbf{q}(t), \mathbf{p}(t), t). \end{aligned} \quad (77)$$

The coordinates $\mathbf{q}$ and $\mathbf{p}$, respectively, each have k components.

F.1 Hamiltonian Vector Fields and Symplectic Forms

We define a Hamiltonian vector field X_H in one time step $t = \tau$ as

$$X_H(\mathbf{q}, \mathbf{p}, \tau) := \begin{pmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{pmatrix}(\mathbf{q}, \mathbf{p}, \tau) = \begin{pmatrix} \frac{\partial H}{\partial \mathbf{p}} \\ -\frac{\partial H}{\partial \mathbf{q}} \end{pmatrix}(\mathbf{q}, \mathbf{p}, \tau). \quad (78)$$

A symplectic form ω is a closed non-degenerate 2-form [16] on a manifold M , defining a specific *isomorphism* between covectors and vectors on M , i.e., between the cotangent bundle⁷ of the manifold and its tangent bundle. The symplectic form ω turns the manifold M into a *symplectic manifold* (M, ω) . In contrast to *Riemannian manifolds* $(M, \mathbf{g})$, where such an isomorphism is given by a Riemannian metric $\mathbf{g}$ (i.e., the musical isomorphisms $\sharp, \flat$, which in coordinate components correspond to raising and lowering of indices), on a symplectic manifold (M, ω) the different structure of ω determines this isomorphism: For each covector (1-form) field dH , for a given scalar field H , there is a unique corresponding vector field X_H , defined as follows.

We first fix a time $t = \tau$, or alternatively assume $\partial H / \partial t = 0$. Given a symplectic form ω , the isomorphism between the cotangent bundle and the tangent bundle is defined by the isomorphism between a covector dH , in the cotangent space at each point x , and a vector X_H , in the corresponding tangent space at the same point x , by requiring that the relationship between dH and X_H is (using our sign convention)

$$dH = \iota_{X_H} \omega. \quad (79)$$

The operator ι denotes the *interior product*, which in the above performs a tensor contraction between the vector field X_H and the 2-form ω . We can also expand Eq. 79 using an arbitrary vector argument Y and write

$$dH(Y) = \omega(X_H, Y). \quad (80)$$

In *canonical coordinates* (q^i, p_i) we can write $X_H = \frac{\partial H}{\partial p_i} \frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i} \frac{\partial}{\partial p_i}$ with the canonical symplectic matrix Ω (Eq. 30; note sign convention),

$$X_H = (dH \Omega)^T, \quad dH = \frac{\partial H}{\partial q^i} dq^i + \frac{\partial H}{\partial p_i} dp_i = \left(\frac{\partial H}{\partial \mathbf{q}} \quad \frac{\partial H}{\partial \mathbf{p}} \right). \quad (81)$$

Because $dH(Y) = (X_H)^T \Omega^{-1} Y$, the matrix $\Omega^{-1} = \Omega^T$ gives the symplectic form ω in canonical coordinates (q^i, p_i) , with ω defined by

$$\omega = dq^i \wedge dp_i = dq^i \otimes dp_i - dp_i \otimes dq^i. \quad (82)$$

Here, $\wedge$ denotes the (anti-symmetric) wedge (or exterior) product of differential forms, $\otimes$ is the tensor product. We can also obtain the same as $\omega = -dp_i \wedge dq^i$, due to the anti-symmetry of the wedge product.

F.2 Derivative of Hamiltonian Along Hamiltonian Trajectory

Interestingly, simply by the anti-symmetry of the symplectic 2-form ω , again for fixed time $t = \tau$, or for $\partial H / \partial t = 0$, we immediately get

$$dH(X_H) = \omega(X_H, X_H) = 0. \quad (83)$$

Thus, for Hamiltonians with $\partial H / \partial t = 0$, we get that H is *constant along any integral curve* of the Hamiltonian vector field X_H , i.e., for time-independent H it is constant along any *Hamiltonian trajectory*.

⁷This is *not* the $2k$ -dimensional cotangent bundle of the k -dimensional configuration space that is the symplectic manifold itself (i.e., the Hamiltonian phase space), but the $4k$ -dimensional cotangent bundle of the $2k$ -dimensional cotangent bundle of the k -dimensional configuration space.

More generally, however, as for unsteady input fields $\mathbf{v}(x, t)$ in our framework, we have $\partial H / \partial t \neq 0$. However, for this case we can obtain

$$dH(\bar{X}_H) = \omega(X_H, X_H) + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t}. \quad (84)$$

On the left-hand side, we have $\bar{X}_H$, as given by Eq. 19, which is the $(2k+1)$ -dimensional equivalent of X_H , as given by Eq. 78. We can convert between the two fields X_H and $\bar{X}_H$, respectively, by simply appending a $(2k+1)$ 'th component of 1 (time advancing “uniformly”), to go from Eq. 78 to Eq. 19, or by choosing a fixed time $t = \tau$ and dropping the last component, to go from Eq. 19 to Eq. 78.

F.3 The Hamiltonian Vector Field from the Lagrangian

If we do not want Eq. 25b to explicitly depend on $\dot{\mathbf{q}}$ on the right-hand side, we can simply insert Eq. 25a into Eq. 25b. This gives us the derivatives $(\dot{\mathbf{q}}, \dot{\mathbf{p}})$ in a form that on the right-hand side now only depends on the coordinates $(\mathbf{q}, \mathbf{p})$, and thus Hamilton's equations are

$$\dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t) = L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}^{-1}(t) \mathbf{p} - L_{\dot{\mathbf{q}}\mathbf{q}}^{-1} L_{\dot{\mathbf{q}}\mathbf{q}}(t) \mathbf{q} + L_{\dot{\mathbf{q}}\mathbf{q}}^{-1} \bar{\mathbf{b}}_l(t), \quad (85a)$$

$$\dot{\mathbf{p}}(\mathbf{q}, \mathbf{p}, t) = (L_{\mathbf{q}\mathbf{q}} - L_{\mathbf{q}\dot{\mathbf{q}}} L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}^{-1} L_{\dot{\mathbf{q}}\mathbf{q}})(t) \mathbf{q} + L_{\mathbf{q}\dot{\mathbf{q}}} L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}^{-1}(t) \mathbf{p} + L_{\mathbf{q}\mathbf{q}} L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}^{-1} \bar{\mathbf{b}}_l(t) - \bar{\mathbf{b}}_h(t). \quad (85b)$$

We can now write the above defining a *Hamiltonian vector field* X_H as

$$X_H(\mathbf{q}, \mathbf{p}, t) = \begin{pmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{pmatrix}(\mathbf{q}, \mathbf{p}, t). \quad (86)$$

In our framework, because our Lagrangians L are quadratic, we can see that the Hamiltonian vector field X_H is determined by an affine map

$$X_H(\mathbf{q}, \mathbf{p}, t) = \mathbf{X}(t) \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} + X_H(0, 0, t). \quad (87)$$

The matrix $\mathbf{X}(t)$ and the affine part $X_H(0, 0, t)$, respectively, can be directly read off of Eqs. 85a, 85b. This gives Eq. 27 and Eq. 28.

F.4 Hamiltonian Stream Lines

As for any linear (affine) vector field, for a Hamiltonian vector field given by Eq. 26 we know that, for any time t , we have

$$\nabla X_H(t) = \mathbf{X}(t). \quad (88)$$

For the special case of a time-independent Hamiltonian H , or a fixed time $t = \tau$, and thus fixed matrix $\mathbf{X}(\tau)$, all “stream line” integral curves of the Hamiltonian vector field are therefore obtained as

$$\tilde{t} \mapsto \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix}(\tilde{t}) = e^{(\tilde{t}-\tau) \cdot \mathbf{X}(\tau)} \cdot \begin{pmatrix} \mathbf{q} - \mathbf{q}_0 \\ \mathbf{p} - \mathbf{p}_0 \end{pmatrix}(\tau) + \begin{pmatrix} \mathbf{q}_0 \\ \mathbf{p}_0 \end{pmatrix}(\tau). \quad (89)$$

Here, $(\mathbf{q}, \mathbf{p})(\tau)$ is the coordinate of an arbitrary start point, at time $t = \tau$, on one particular integral curve of the Hamiltonian vector field X_H , and we have used the matrix exponential. For the above equation, we have to compute the coordinates $(\mathbf{q}_0, \mathbf{p}_0)(\tau)$ of a critical point of the field X_H , for fixed time $t = \tau$. Assuming $\mathbf{X}(\tau)$ is invertible, there will be exactly one isolated critical point with coordinates

$$\begin{pmatrix} \mathbf{q}_0 \\ \mathbf{p}_0 \end{pmatrix}(\tau) = -\mathbf{X}^{-1}(\tau) \begin{pmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{pmatrix}(0, 0, \tau). \quad (90)$$

Even for time-dependent Hamiltonians H , as in our framework for unsteady input fields $\mathbf{v}(x, t)$, we can use the equation above to determine all “stream lines” of the field X_H for fixed, arbitrary time $t = \tau$. We can furthermore also investigate the qualitative behavior of the vector field X_H , by looking at the different possible Jordan normal forms of the matrix $\mathbf{X}$. Moreover, because the matrix $\mathbf{X}$ fulfills the *symplectic condition* $\mathbf{X}^T \Omega + \Omega \mathbf{X} = 0$, making the matrix $\mathbf{X}$ an infinitesimally symplectic matrix, we have the corresponding *symplectic matrix*⁸

$$e^{t\mathbf{X}}. \quad (91)$$

⁸The symplectic matrices form a Lie group, and for a given symplectic matrix $e^{t\mathbf{X}}$, the matrix $\mathbf{X}$ is an element of the corresponding Lie algebra.

F.5 Hamiltonian Trajectories (Path Lines)

Hamiltonian trajectories, i.e., path lines, of a time-dependent Hamiltonian vector field $\bar{X}_H$ given by Eq. 19 can be obtained as initial value problem solutions

$$t \mapsto \begin{pmatrix} \mathbf{q}(t) \\ \mathbf{p}(t) \\ t \end{pmatrix} = \begin{pmatrix} \mathbf{q}(t_0) \\ \mathbf{p}(t_0) \\ t_0 \end{pmatrix} + \int_{t_0}^t \bar{X}_H(\mathbf{q}(\tau), \mathbf{p}(\tau), \tau) d\tau. \quad (92)$$

F.6 Numerical Integration of Hamiltonian Trajectories

In order to integrate a Hamiltonian vector field X_H , *symplectic integrators* are often used because they preserve the symplectic structure of the Hamiltonian phase space.⁹ However, they are often implicit integrators. If the Hamiltonian is separable in the variables $(\mathbf{q}, \mathbf{p})$, explicit integrators can be used. Unfortunately, in our framework this is in general not the case, because in Eqs. 26, 29, Hamilton's equations are only separable when $\mathbf{A} = 0$.¹⁰ However, the fact that our fields X_H are linear (affine) also greatly simplifies symplectic integration.

F.6.1 Second-order symplectic integrator

A simple second-order (1-stage)¹¹ symplectic integrator that can exploit the linear structure of fields X_H given by Eq. 26 is the *implicit midpoint rule*. Abbreviating $\mathbf{x}(t) := (\mathbf{q}, \mathbf{p})^T(t)$, $\mathbf{r}(t) := (\dot{\mathbf{q}}, \dot{\mathbf{p}})^T(0, 0, t)$, the *implicit* update rule to advance from time t_j to time t_{j+1} , with step size $h = t_{j+1} - t_j$, and midpoint time $t_m = t_j + \frac{h}{2}$, is given by

$$\begin{aligned} \mathbf{x}(t_{j+1}) &= \mathbf{x}(t_j) + h X_H \left(\frac{\mathbf{x}(t_j) + \mathbf{x}(t_{j+1})}{2}, t_m \right), \\ &= \mathbf{x}(t_j) + h \mathbf{X}(t_m) \frac{\mathbf{x}(t_j) + \mathbf{x}(t_{j+1})}{2} + h \mathbf{r}(t_m). \end{aligned} \quad (93)$$

Now, linearity allows solving for $\mathbf{x}(t_{j+1})$, giving the *explicit* update

$$\mathbf{x}(t_{j+1}) = \left(\mathbf{I} - \frac{h}{2} \mathbf{X}(t_m) \right)^{-1} \left(\left(\mathbf{I} + \frac{h}{2} \mathbf{X}(t_m) \right) \mathbf{x}(t_j) + h \mathbf{r}(t_m) \right). \quad (94)$$

F.6.2 Fourth-order symplectic integrator

For a fourth-order (2-stage) Gauss-Legendre scheme, we first have to compute two specific intermediate time points in the interval $[t_j, t_{j+1}]$:

$$t_1 = t_j + h \left(\frac{1}{2} - \frac{\sqrt{3}}{6} \right), \quad t_2 = t_j + h \left(\frac{1}{2} + \frac{\sqrt{3}}{6} \right). \quad (95)$$

We then solve for intermediate $2k$ -vectors $\mathbf{x}_{1|2}$ in the $4k \times 4k$ system

$$\begin{pmatrix} \mathbf{I} - \frac{h}{4} \mathbf{X}(t_1) & -h \left(\frac{1}{4} - \frac{\sqrt{3}}{6} \right) \mathbf{X}(t_2) \\ -h \left(\frac{1}{4} + \frac{\sqrt{3}}{6} \right) \mathbf{X}(t_1) & \mathbf{I} - \frac{h}{4} \mathbf{X}(t_2) \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix}. \quad (96)$$

The right-hand side $2k$ -vectors $\mathbf{b}_{1|2}$ above are given by

$$\begin{aligned} \mathbf{b}_1 &= \mathbf{x}(t_j) + \frac{h}{4} \mathbf{r}(t_1) + h \left(\frac{1}{4} - \frac{\sqrt{3}}{6} \right) \mathbf{r}(t_2), \\ \mathbf{b}_2 &= \mathbf{x}(t_j) + \frac{h}{4} \mathbf{r}(t_2) + h \left(\frac{1}{4} + \frac{\sqrt{3}}{6} \right) \mathbf{r}(t_1). \end{aligned} \quad (97)$$

The final update is then given by

$$\mathbf{x}(t_{j+1}) = \mathbf{x}(t_j) + \frac{h}{2} \left(\mathbf{X}(t_1) \mathbf{x}_1 + \mathbf{r}(t_1) + \mathbf{X}(t_2) \mathbf{x}_2 + \mathbf{r}(t_2) \right). \quad (98)$$

⁹This means that the symplectic form ω is preserved under the *flow* of the Hamiltonian vector field: $\mathcal{L}_{X_H} \omega = 0$, with $\mathcal{L}$ denoting the Lie derivative.

¹⁰We have that $\mathbf{A} = 0$ if and only if $L_{\mathbf{q}\dot{\mathbf{q}}} = L_{\dot{\mathbf{q}}\mathbf{q}}^T = 0$ (cf. Eq. 27), because $L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}$ is invertible. In general, our Lagrangians are not separable (cf. Eq. 160).

¹¹The midpoint rule is the simplest member of the Gauss-Legendre family of symplectic integrators. Each s -stage method of this family is $2s$ -order accurate.

F.6.3 Tracking trajectory integration error over time

We can make use of the properties of a time-dependent Hamiltonian function H in order to track the accuracy of integrating a Hamiltonian trajectory over time. In order to do this, we can change the non-autonomous (i.e., explicitly time-dependent) system in the Hamiltonian phase space $(\mathbf{q}, \mathbf{p}, t)$, with a Hamiltonian $H(\mathbf{q}, \mathbf{p}, t)$, to an autonomous system in a $2(k+1)$ -dimensional extended phase space $(\mathbf{q}, \mathbf{p}, q_t, p_t)$, where we can define an extended, *time-independent* Hamiltonian $\bar{H}(\mathbf{q}, \mathbf{p}, q_t, p_t)$. In this setup, q_t represents a generalized coordinate for time (but time is not explicit), and p_t represents the corresponding generalized momentum. In the extended phase space $(\mathbf{q}, \mathbf{p}, q_t, p_t)$, we can now define the extended Hamiltonian as

$$\bar{H}(\mathbf{q}, \mathbf{p}, q_t, p_t) = H(\mathbf{q}, \mathbf{p}, t) + p_t. \quad (99)$$

This definition is such that by evaluating Hamilton's equations for the extended Hamiltonian $\bar{H}$ explicitly via the partial derivatives of the function $\bar{H}$, we get the corresponding Hamiltonian vector components

$$\begin{aligned} \dot{q}_t(\mathbf{q}, \mathbf{p}, q_t, p_t) &= \frac{\partial \bar{H}}{\partial p_t} = 1, \\ \dot{p}_t(\mathbf{q}, \mathbf{p}, q_t, p_t) &= -\frac{\partial \bar{H}}{\partial q_t} = -\frac{\partial H(\mathbf{q}, \mathbf{p}, t)}{\partial t} \Big|_{t=q_t}. \end{aligned} \quad (100)$$

To integrate a trajectory $t \mapsto (\mathbf{q}, \mathbf{p}, q_t, p_t)(t) = (\mathbf{x}, q_t, p_t)(t)$ from some start time $t = t_0$, we start with the coordinate $q_t(t_0) = t_0$, and the momentum $p_t(t_0) = -H(\mathbf{q}(t_0), \mathbf{p}(t_0), t_0)$. In this way, along the (accurate) trajectory we must have the *constant* extended Hamiltonian

$$\bar{H}(\mathbf{q}(t), \mathbf{p}(t), q_t(t), p_t(t)) = 0. \quad (101)$$

We simply integrate the additional vector field components $(\dot{q}_t, \dot{p}_t)(t)$ along the usual components $(\dot{\mathbf{q}}, \dot{\mathbf{p}})(t)$. The integrated time coordinate q_t is immediately known to be $q_t(t_{j+1}) = q_t(t_j) + h = t_{j+1}$, so we do not actually need to compute it. When the second-order (1-stage) midpoint rule is used, the update of the integrated momentum p_t is

$$p_t(t_{j+1}) = p_t(t_j) - h \frac{\partial H(\mathbf{x}(t_m), t)}{\partial t} \Big|_{t=t_m}. \quad (102)$$

We can evaluate this by first computing $\mathbf{x}(t_{j+1})$, and then the midpoint

$$\mathbf{x}(t_m) = \frac{\mathbf{x}(t_j) + \mathbf{x}(t_{j+1})}{2}. \quad (103)$$

In order to estimate whether the integration error is too large, we can use the actual value of the extended Hamiltonian $\bar{H}$, i.e., we compute

$$\begin{aligned} \delta(t_{j+1}) &= \bar{H}(\mathbf{x}(t_{j+1}), q_t(t_{j+1}), p_t(t_{j+1})), \\ &= H(\mathbf{x}(t_{j+1}), t_{j+1}) + p_t(t_{j+1}). \end{aligned} \quad (104)$$

For exact integration, we would have $\delta = 0$ at every step, as given by Eq. 101, which we defined for integrating a trajectory with a start value of $\bar{H}(t_0) = 0$, i.e., an integration constant $p_t(t_0) = -H(t_0)$. When the error $|\delta|$ becomes too large, we can switch from second-order integration to fourth-order, and/or reduce the step size h appropriately.

When the fourth-order (2-stage) symplectic integrator above is used, we have to update the p_t coordinate in the same way, by first computing

$$\dot{H}_1 = \frac{\partial H(\mathbf{x}_1, t)}{\partial t} \Big|_{t=t_1}, \quad \dot{H}_2 = \frac{\partial H(\mathbf{x}_2, t)}{\partial t} \Big|_{t=t_2}. \quad (105)$$

The final update for p_t in the next step is then given by

$$p_t(t_{j+1}) = p_t(t_j) - \frac{h}{2} (\dot{H}_1 + \dot{H}_2). \quad (106)$$

We can again perform the same accuracy check using $|\delta|$ (Eq. 104).

Finally, we note that the autonomous, extended phase space formulation above, and the error δ , can in fact also be used to perform a correction in each numerical integration step by projecting the numerical position back onto the actual trajectory, preserving symplecticity [17].

G GENERAL FORM OF QUADRATIC HAMILTONIANS

We derive the general form of the Hamiltonian function H when the Lagrangian L is given in the general form of Eq. 11. From Eq. 67, and by using Eq. 23, we can write the general form of the energy E_L (which defines the Hamiltonian H via Eq. 17) as

$$\begin{aligned} E_L(\mathbf{q}, \dot{\mathbf{q}}, t) &= (L_{\dot{\mathbf{q}}\mathbf{q}}(t) \mathbf{q} + L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}(t) \dot{\mathbf{q}} - \bar{\mathbf{b}}_l(t))^T \dot{\mathbf{q}} - L, \\ &= \mathbf{q}^T L_{\mathbf{q}\dot{\mathbf{q}}}(t) \dot{\mathbf{q}} + \dot{\mathbf{q}}^T L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}(t) \dot{\mathbf{q}} - \bar{\mathbf{b}}_l^T(t) \dot{\mathbf{q}} - L. \end{aligned} \quad (107)$$

Inserting L from Eq. 60, we obtain the expansion

$$\begin{aligned} E_L(\mathbf{q}, \dot{\mathbf{q}}, t) &= \mathbf{q}^T L_{\mathbf{q}\dot{\mathbf{q}}}(t) \dot{\mathbf{q}} + \dot{\mathbf{q}}^T L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}(t) \dot{\mathbf{q}} - \bar{\mathbf{b}}_l^T(t) \dot{\mathbf{q}} \\ &\quad - \frac{1}{2} \begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix}^T \mathbf{H}_L(t) \begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix} + \begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix}^T \mathbf{A}^T \mathbf{b}(t) - \frac{1}{2} \mathbf{b}^T \mathbf{b}(t). \end{aligned} \quad (108)$$

Now, however, we can furthermore insert the expansion

$$\begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix}^T \mathbf{H}_L(t) \begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix} = \mathbf{q}^T L_{\mathbf{q}\mathbf{q}}(t) \mathbf{q} + 2 \mathbf{q}^T L_{\mathbf{q}\dot{\mathbf{q}}}(t) \dot{\mathbf{q}} + \dot{\mathbf{q}}^T L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}(t) \dot{\mathbf{q}}. \quad (109)$$

Inserting this expansion into Eq. 108, the energy E_L simplifies to

$$\begin{aligned} E_L(\mathbf{q}, \dot{\mathbf{q}}, t) &= \frac{1}{2} \dot{\mathbf{q}}^T L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}(t) \dot{\mathbf{q}} - \bar{\mathbf{b}}_l^T(t) \dot{\mathbf{q}} \\ &\quad - \frac{1}{2} \mathbf{q}^T L_{\mathbf{q}\mathbf{q}}(t) \mathbf{q} + (\mathbf{q}^T, \dot{\mathbf{q}}^T) \mathbf{A}^T \mathbf{b}(t) - \frac{1}{2} \mathbf{b}^T \mathbf{b}(t). \end{aligned} \quad (110)$$

Thus, the energy $E_L(\mathbf{q}, \dot{\mathbf{q}}, t)$ for any Lagrangian L of the general form given by Eq. 11 is *separable*, with respect to the coordinates $\mathbf{q}$ and $\dot{\mathbf{q}}$.

Given the energy $E_L(\mathbf{q}, \dot{\mathbf{q}}, t)$, the Hamiltonian $H(\mathbf{q}, \mathbf{p}, t)$ is defined by Eq. 17. It is therefore possible to compute $H(\mathbf{q}, \mathbf{p}, t)$ from Eq. 110 together with the coordinate transformation $\dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t)$ given by Eq. 24:

$$\begin{aligned} H(\mathbf{q}, \mathbf{p}, t) &= \frac{1}{2} \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t)^T L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}(t) \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t) - \bar{\mathbf{b}}_l^T(t) \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t) \\ &\quad - \frac{1}{2} \mathbf{q}^T L_{\mathbf{q}\mathbf{q}}(t) \mathbf{q} + (\mathbf{q}^T, \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t)^T) \mathbf{A}^T \mathbf{b}(t) - \frac{1}{2} \mathbf{b}^T \mathbf{b}(t). \end{aligned} \quad (111)$$

However, this expression is not fully simplified in $(\mathbf{q}, \mathbf{p})$, and computing partial derivatives would require extensive use of the chain rule due to the coordinate transformation $\dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t)$. We can nevertheless obtain the final simplified expression for $H(\mathbf{q}, \mathbf{p}, t)$ in Eq. 32 without explicitly taking derivatives. First, similar to Eq. 63, we define the general structure of the Hessian matrix of the Hamiltonian H as

$$\mathbf{H}_H(t) = \begin{pmatrix} \frac{\partial^2 H}{\partial q^i \partial q^j} & \frac{\partial^2 H}{\partial q^i \partial p_j} \\ \frac{\partial^2 H}{\partial p_i \partial q^j} & \frac{\partial^2 H}{\partial p_i \partial p_j} \end{pmatrix} =: \begin{pmatrix} \mathbf{H}_{\mathbf{q}\mathbf{q}}(t) & \mathbf{H}_{\mathbf{q}\mathbf{p}}(t) \\ \mathbf{H}_{\mathbf{p}\mathbf{q}}(t) & \mathbf{H}_{\mathbf{p}\mathbf{p}}(t) \end{pmatrix}. \quad (112)$$

Now, using the relationship in Eq. 31 gives Eq. 32, and we therefore already know all matrix blocks of $\mathbf{H}_H(t)$ above *without actually computing partial derivatives*. The Hamiltonian can then be computed simply by inserting the matrix blocks of $\mathbf{H}_H(t)$ into Eq. 111 and re-arranging. This gives the final expression for the Hamiltonian $H(\mathbf{q}, \mathbf{p}, t)$ in Eq. 32.

H OBSERVER CURVES DETERMINED POINT-WISE

We can ask a more general question than Theorem 1 and Theorem 2, which will turn out to contain both as special cases. The question is: When is a curve $t \mapsto (\mathbf{q}, \dot{\mathbf{q}}, t)(t)$, or, equivalently, a curve $t \mapsto (\mathbf{q}, \mathbf{p}, t)(t)$, which was computed point-wise via Eq. 34 to connect unique global minima (assuming the Hessian $\mathbf{H}_L$ is invertible) in the Lagrangian phase space $(\mathbf{q}, \dot{\mathbf{q}}, t)$, a curve that fulfills Eq. 45 and the Euler-Lagrange equation, or, equivalently, Hamilton's equations?

The following theorem answers this question, where we obtain

Theorem 3. *Assuming the Hessian matrix $\mathbf{H}_L$ of a Lagrangian L is invertible, a curve $t \mapsto (\mathbf{q}, \mathbf{p}, t)(t)$ with $\mathbf{p}(t) = 0$, computed point-wise via Eq. 34, is a Hamiltonian trajectory if and only if it fulfills*

$$\begin{aligned} \frac{\partial}{\partial t} H(\mathbf{q}(t), 0, t) &= \frac{d}{dt} H(\mathbf{q}(t), 0, t), \\ &= -\frac{d}{dt} L(\mathbf{q}(t), \dot{\mathbf{q}}(\mathbf{q}(t), 0, t), t). \end{aligned} \quad (113)$$

Proof. First, it is important to note that the *left-hand* side of the equation above gives the directional derivative $dH(\bar{X}_H)$ in the direction of the Hamiltonian trajectory at the point $(\mathbf{q}(t), 0, t)$, although neither $\bar{X}_H$ nor the directional derivative are computed explicitly. Likewise, the *right-hand* side of the equation above denotes the derivative of the function H restricted to the curve $t \mapsto (\mathbf{q}, \mathbf{p}, t)(t)$, where $\mathbf{p}(t) = 0$ is always fulfilled (and therefore also $H = -L$), which, in general, is not the same as the Hamiltonian trajectory with tangent $\bar{X}_H$ at the same point. This derivative also need not be computed explicitly as a directional derivative along the curve, because it simply is the derivative of a 1D scalar function $H(t) = -L(\mathbf{q}(t), \dot{\mathbf{q}}(\mathbf{q}(t), 0, t), t)$ and can easily be computed as such. Now, in order for the Hamiltonian trajectory to agree with the curve $t \mapsto (\mathbf{q}, 0, t)(t)$, because of Eq. 34 everywhere on the Hamiltonian trajectory we would have to fulfill the condition

$$\frac{d}{dt} \frac{\partial L}{\partial \mathbf{q}} = d \left(\frac{\partial L}{\partial \mathbf{q}} \right) (\bar{X}_H) = 0. \quad (114)$$

It is crucial that here, in contrast to Eq. 113, d/dt denotes the directional derivative in the direction $\bar{X}_H$. A general Hamiltonian trajectory will not fulfill this condition. Now, assuming that the Hessian matrix $\mathbf{H}_L$ is invertible, for each time t there will be one unique minimum of the scalar function L , for which Eq. 34 gives the corresponding coordinate $(\mathbf{q}, 0, t)$. For Eq. 114 to be fulfilled, because the minimum of L must occur where $\partial L / \partial \mathbf{q} = \partial L / \partial \dot{\mathbf{q}} = 0$, if Eq. 113 is fulfilled, we also know that Eq. 114 will be fulfilled. Conversely, if Eq. 113 is not fulfilled, Eq. 114 cannot be fulfilled because then $\partial L / \partial \mathbf{q} \neq 0$. This proves the theorem, because this immediately means that the two curves will agree if and only if Eq. 113 is fulfilled. $\square$

Corollary 4. *If Eq. 113 is fulfilled for a curve $t \mapsto (\mathbf{q}, 0, t)(t)$ computed via Eq. 34, because the Hamiltonian trajectory will agree with the curve $t \mapsto (\mathbf{q}, 0, t)(t)$, and thus the curve $t \mapsto (\mathbf{q}, \dot{\mathbf{q}}(\mathbf{q}(t), 0, t), t)(t)$, the tangent condition given in Eq. 45 will be fulfilled.*

The corollary is an immediate consequence of Theorem 3 because if the two curves agree, the tangent vector field of the curve $t \mapsto (\mathbf{q}, 0, t)(t)$ must agree with the coordinates $\dot{\mathbf{q}} = \dot{\mathbf{q}}(\mathbf{q}(t), 0, t)$ along the curve, because the Hamiltonian trajectory fulfills this property. Because the latter is defined via integration, it naturally guarantees Eq. 45.

H.1 Hamiltonian and Partial Time Derivative

From Eq. 32, for the fixed subspace $\mathbf{p} = 0$ we obtain

$$H(\mathbf{q}, 0, t) = \frac{1}{2} \mathbf{q}^T \mathbf{H}_{\mathbf{q}\mathbf{q}}(t) \mathbf{q} - \mathbf{q}^T \dot{\mathbf{p}}(0, 0, t) + \bar{\mathbf{b}}_l^T L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}^{-1} \bar{\mathbf{b}}_l(t) - \frac{1}{2} \mathbf{b}^T \mathbf{b}(t). \quad (115)$$

In the subspace $\mathbf{p} = 0$, the partial time derivative of $H(\mathbf{q}, 0, t)$ thus is

$$\begin{aligned} \frac{\partial H}{\partial t}(\mathbf{q}, 0, t) &= \frac{1}{2} \mathbf{q}^T \frac{\partial \mathbf{H}_{\mathbf{q}\mathbf{q}}}{\partial t}(t) \mathbf{q} - \mathbf{q}^T \frac{\partial \dot{\mathbf{p}}}{\partial t}(0, 0, t) \\ &\quad + \frac{\partial}{\partial t} \left(\bar{\mathbf{b}}_l^T L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}^{-1} \bar{\mathbf{b}}_l - \frac{1}{2} \mathbf{b}^T \mathbf{b} \right)(t). \end{aligned} \quad (116)$$

However, to evaluate Eq. 46, we do not need to know the Hamiltonian H , because we can compute everything from L and its derivatives:

1. We can compute $H(\mathbf{q}, 0, t)$ from the Lagrangian L via

$$H(\mathbf{q}, 0, t) = -L(\mathbf{q}, \dot{\mathbf{q}}(\mathbf{q}, 0, t), t), \quad (117)$$

2. Eq. 116 can be computed using $\partial H_{\mathbf{q}\mathbf{q}} / \partial t = -\partial \mathbf{C} / \partial t$ from the matrix $\mathbf{C}$ (Eq. 27), and the vector $\partial \dot{\mathbf{p}}(0, 0, t) / \partial t$ (Eqs. 26, 28).

I DISCRETIZED OBSERVER CURVES

We represent a reference frame, i.e., a time-dependent observer, as the entire curve $t \mapsto \mathbf{q}(t)$, but discretized in time with T time steps. The corresponding discrete samples are packed into one large $k \times T$ vector

$\bar{\mathbf{q}}$ for all time steps $t \in [t_a, t_b]$, i.e., for time steps $\{t_a, t_a + \Delta t, t_a + 2\Delta t, \dots, t_b\}$. Thus, the observer in discretized form is a vector

$$\bar{\mathbf{q}} = \begin{pmatrix} \mathbf{q}(t_a) \\ \mathbf{q}(t_a + \Delta t) \\ \mathbf{q}(t_a + 2\Delta t) \\ \vdots \\ \mathbf{q}(t_b) \end{pmatrix}. \quad (118)$$

Euclidean observers. For observers modeled as exact Killing fields, the parameters $\mathbf{q}(\tau)$ for time $t = \tau$ are given by k components q^i ($k = 3$ for the domains $\mathbb{R}^2$ and $\mathbb{S}^2$; $k = 6$ for the domain $\mathbb{R}^3$), i.e.,

$$\mathbf{q}(\tau) = (q^i)_{i=1}^k(\tau). \quad (119)$$

From the parameters $q^i(t)$, the Killing field $\mathbf{u}$ is obtained as $\mathbf{u}(x, t) = \sum_{i=1}^k q^i(t) \mathbf{e}_i(x, t)$, with the vector fields $\mathbf{e}_i$ a basis of k linearly independent Killing fields given on the same domain as the input flow $\mathbf{v}$.

General observer fields. For observers modeled as *approximate* Killing fields $\mathbf{u}$, we simply store discretized general vector fields. The parameters $\mathbf{q}(\tau)$ for time $t = \tau$ are given by $k = n \cdot \tilde{N}$ coefficients, where $\tilde{N}$ is the number of spatial samples for one time step, and $n = 2$ (2D domains, i.e., 2D vector fields), or $n = 3$ (3D domains, i.e., 3D vector fields). For example, for a 2D domain with $N = I \cdot J$ grid points with coordinates (x_i, y_j) , and $\tilde{N} = N$, i.e., n coefficients per grid point where we use $k = n \cdot N$ spatial box basis functions $\mathbf{e}_i$,

$$\mathbf{q}(\tau) = (q^i)_{i=1}^k(\tau) = (u^1(x_1, y_1, \tau), u^2(x_1, y_1, \tau), \dots, u^1(x_I, y_J, \tau), u^2(x_I, y_J, \tau)). \quad (120)$$

I.1 Lagrangian Phase Space

Where needed, we compute time derivatives $\dot{\mathbf{q}}(t) = (d/dt) \mathbf{q}(t)$ via the linear map

$$\begin{pmatrix} \mathbf{q}(t_a) \\ \dot{\mathbf{q}}(t_a) \\ \vdots \\ \mathbf{q}(t_b) \\ \dot{\mathbf{q}}(t_b) \end{pmatrix} = \bar{\mathbf{D}} \bar{\mathbf{q}}. \quad (121)$$

Here, the linear operator $\bar{\mathbf{D}}$ computes first derivatives by using finite differencing of the function $t \mapsto \mathbf{q}(t)$ in time.

We also define the following extraction of $(\mathbf{q}, \dot{\mathbf{q}})(\tau)$ for one specific time $t = \tau$:

$$\begin{pmatrix} \mathbf{q}(\tau) \\ \dot{\mathbf{q}}(\tau) \end{pmatrix} = \bar{\mathbf{T}}(\tau) \bar{\mathbf{D}} \bar{\mathbf{q}}. \quad (122)$$

The operator to extract $(\mathbf{q}, \dot{\mathbf{q}})(\tau)$ at one time τ is

$$\bar{\mathbf{T}}(\tau) = \left(\begin{array}{c|c|c} \cdots & \mathbf{I} & \mathbf{0} \\ \cdots & \mathbf{0} & \mathbf{I} \end{array} \right). \quad (123)$$

I.2 Hamiltonian Phase Space

A curve $t \mapsto (\mathbf{q}, \mathbf{p})(t)$ through the Hamiltonian phase space is stored in discretized form as a $2k \times T$ vector

$$\begin{pmatrix} \mathbf{q}(t_a) \\ \mathbf{p}(t_a) \\ \vdots \\ \mathbf{q}(t_b) \\ \mathbf{p}(t_b) \end{pmatrix}. \quad (124)$$

For optimal observers, this vector stores a discretized version of an integral curve of the corresponding Hamiltonian vector field X_H .

The vector field X_H along this integral curve is given by

$$\begin{pmatrix} \dot{\mathbf{q}}(t_a) \\ \dot{\mathbf{p}}(t_a) \\ \vdots \\ \dot{\mathbf{q}}(t_b) \\ \dot{\mathbf{p}}(t_b) \end{pmatrix}. \quad (125)$$

J EULER-LAGRANGE OPERATOR, FUNCTIONAL DERIVATIVE

Given a curve $t \mapsto \mathbf{q}(t)$, for reference we first repeat the action (integral) S defined in Eq. 13, given by

$$S[\mathbf{q}] = \int_{t_a}^{t_b} L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t) dt.$$

J.1 First Variation

Given the action S , corresponding to integrating the Lagrangian L over the time interval $[t_a, t_b]$, we define the corresponding *functional differential* of the action S with respect to the function $\mathbf{q}$ as

$$\delta S[\mathbf{q}, \boldsymbol{\eta}] = \frac{d}{d\epsilon} \left[S[\mathbf{q} + \epsilon \boldsymbol{\eta}] \right] = \lim_{\epsilon \rightarrow 0} \frac{S[\mathbf{q} + \epsilon \boldsymbol{\eta}] - S[\mathbf{q}]}{\epsilon}. \quad (126)$$

This differential is defined using a *variation* $\boldsymbol{\eta}(t)$, which is a smooth function defined over the same interval as the function $\mathbf{q}(t)$, often with pre-scribed boundary conditions (see below). This variation is often also written as a function $\delta \mathbf{q}$, i.e., we would often also write $\delta \mathbf{q} = \boldsymbol{\eta}$.

We can differentiate under the integral to explicitly compute the limit above as (for brevity, we write δS instead of $\delta S[\mathbf{q}, \boldsymbol{\eta}]$ below)

$$\begin{aligned} \delta S &= \int_{t_a}^{t_b} \left(\frac{\partial L}{\partial(\mathbf{q} + \epsilon \boldsymbol{\eta})} \cdot \frac{d(\mathbf{q} + \epsilon \boldsymbol{\eta})}{d\epsilon} + \frac{\partial L}{\partial(\dot{\mathbf{q}} + \epsilon \dot{\boldsymbol{\eta}})} \cdot \frac{d(\dot{\mathbf{q}} + \epsilon \dot{\boldsymbol{\eta}})}{d\epsilon} \right) (t) dt, \\ &= \int_{t_a}^{t_b} \left(\frac{\partial L}{\partial(\mathbf{q} + \epsilon \boldsymbol{\eta})} \cdot \boldsymbol{\eta}(t) + \frac{\partial L}{\partial(\dot{\mathbf{q}} + \epsilon \dot{\boldsymbol{\eta}})} \cdot \dot{\boldsymbol{\eta}}(t) \right) dt. \end{aligned} \quad (127)$$

It is important to note that for vector-valued $\mathbf{q}$, i.e., $\mathbf{q} \in \mathbb{R}^k$ with $k > 1$, the vector-valued expressions above must be used in “component dot products,” which we can also denote by tensor contractions using the Einstein summation convention and writing covariant and contravariant indices, respectively. That is, an expression like $\frac{\partial L}{\partial \mathbf{q}} \cdot \boldsymbol{\eta}(t)$ must be read

$$\frac{\partial L}{\partial \mathbf{q}} \cdot \boldsymbol{\eta}(t) = \frac{\partial L}{\partial q^i} \eta^i(t) = \sum_i \frac{\partial L}{\partial q^i} \eta^i(t). \quad (128)$$

Moreover, of course expressions such as $\frac{\partial L}{\partial \mathbf{q}}$ and $\frac{\partial L}{\partial \dot{\mathbf{q}}}$, respectively, must all be evaluated at points $(\mathbf{q}(t), \dot{\mathbf{q}}(t), t)$, i.e., they are

$$\frac{\partial L}{\partial \mathbf{q}} = \frac{\partial L}{\partial \mathbf{q}} \Big|_{(\mathbf{q}(t), \dot{\mathbf{q}}(t), t)}, \quad \frac{\partial L}{\partial \dot{\mathbf{q}}} = \frac{\partial L}{\partial \dot{\mathbf{q}}} \Big|_{(\mathbf{q}(t), \dot{\mathbf{q}}(t), t)}. \quad (129)$$

Now, because we evaluate the limit at $\epsilon = 0$, we furthermore get

$$\begin{aligned} \delta S[\mathbf{q}, \boldsymbol{\eta}] &= \int_{t_a}^{t_b} \left(\frac{\partial L}{\partial \mathbf{q}} \cdot \boldsymbol{\eta}(t) + \frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \dot{\boldsymbol{\eta}}(t) \right) dt, \\ &= \int_{t_a}^{t_b} \left(\frac{\partial L}{\partial \mathbf{q}} \cdot \boldsymbol{\eta}(t) + \frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \frac{d}{dt} \boldsymbol{\eta}(t) \right) dt, \\ &= \int_{t_a}^{t_b} \left(\frac{\partial L}{\partial \mathbf{q}} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{q}}} \right) \cdot \boldsymbol{\eta}(t) dt \\ &\quad + \frac{\partial L}{\partial \dot{\mathbf{q}}}(t_b) \cdot \boldsymbol{\eta}(t_b) - \frac{\partial L}{\partial \dot{\mathbf{q}}}(t_a) \cdot \boldsymbol{\eta}(t_a). \end{aligned} \quad (130)$$

In the last row above, we have used integration by parts on the second term in the second row above. Here, we can explicitly see two of the major ways of choosing boundary conditions in order to determine solutions $\mathbf{q}(t)$ where the functional differential vanishes for all possible variations $\boldsymbol{\eta}(t)$ (i.e., curves $t \mapsto \mathbf{q}(t)$ for which the directional derivative in function space vanishes in all possible directions):

- Boundary conditions with known (i.e., pre-scribed) fixed values $\mathbf{q}(t_a)$ and $\mathbf{q}(t_b)$, respectively, leading to the requirement that all variations $\boldsymbol{\eta}(t)$ must vanish at the boundary. That is, we must correspondingly have the constraint $\boldsymbol{\eta}(t_a) = \boldsymbol{\eta}(t_b) = 0$.
- The *natural boundary conditions*, with free boundary values (a priori unknown) $\mathbf{q}(t_a), \mathbf{q}(t_b)$, and not necessarily vanishing $\boldsymbol{\eta}(t_a), \boldsymbol{\eta}(t_b)$. In this case, we require $\frac{\partial L}{\partial \dot{\mathbf{q}}}(t_a) = \frac{\partial L}{\partial \dot{\mathbf{q}}}(t_b) = 0$.

For both of these choices of boundary conditions, the boundary terms in the functional differential above vanish, and we obtain

$$\delta S[\mathbf{q}, \boldsymbol{\eta}] = \int_{t_a}^{t_b} \left(\frac{\partial L}{\partial \mathbf{q}} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{q}}} \right) \cdot \boldsymbol{\eta}(t) dt. \quad (131)$$

The *fundamental lemma of the calculus of variations* now specifies that if the functional differential above vanishes for all compactly supported smooth functions $\boldsymbol{\eta}(t)$ on $t \in (t_a, t_b)$, the term in the parenthesis above must be identically zero, leading to the Euler-Lagrange equation $\boldsymbol{\psi} = 0$ (with $\boldsymbol{\psi}$ in Eq. 14). This is the function space equivalent of Fermat’s theorem in calculus, where the gradient must vanish at extrema.

We can introduce the *functional derivative* $\frac{\delta S}{\delta \mathbf{q}}$ to write Eq. 131 as

$$\delta S[\mathbf{q}, \boldsymbol{\eta}] = \int_{t_a}^{t_b} \frac{\delta S}{\delta \mathbf{q}} \cdot \boldsymbol{\eta}(t) dt. \quad (132)$$

We can elucidate the notation for the functional derivative by writing the variation $\boldsymbol{\eta}$ as $\delta \mathbf{q} = \boldsymbol{\eta}$, and therefore the functional differential as

$$\delta S[\mathbf{q}, \delta \mathbf{q}] = \int_{t_a}^{t_b} \frac{\delta S}{\delta \mathbf{q}} \cdot \delta \mathbf{q}(t) dt. \quad (133)$$

From the above, we have therefore defined the functional derivative as

$$\frac{\delta S}{\delta \mathbf{q}} = \frac{\partial L}{\partial \mathbf{q}} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{q}}}. \quad (134)$$

This is the same as the Euler-Lagrange operator $\boldsymbol{\psi}$ from Eq. 14, i.e.,

$$\boldsymbol{\psi} = \frac{\delta S}{\delta \mathbf{q}}. \quad (135)$$

Furthermore, from the functional differential $\delta S[\mathbf{q}, \boldsymbol{\eta}]$, we can see that the change in δS for a given function $\boldsymbol{\eta}$ is determined by the functional derivative $\delta S/\delta \mathbf{q}$. The “direction” (in function space) of the fastest change is the “direction” (i.e., function) given by $\delta S/\delta \mathbf{q}$ itself. That is, when we choose $\boldsymbol{\eta} = \delta S/\delta \mathbf{q}$, we obtain the fastest change δS as the inner product (of functions) of the same function with itself, given by

$$\delta S \left[\mathbf{q}, \frac{\delta S}{\delta \mathbf{q}} \right] = \int_{t_a}^{t_b} \frac{\delta S}{\delta \mathbf{q}} \cdot \frac{\delta S}{\delta \mathbf{q}}(t) dt. \quad (136)$$

Therefore, we can interpret the functional derivative $\delta S/\delta \mathbf{q}$ as an *infinite-dimensional gradient* in function space, which we can use in a gradient descent procedure to iteratively approach a point (in function space) where the gradient $\delta S/\delta \mathbf{q}$ vanishes, i.e., where $\delta S/\delta \mathbf{q} = \boldsymbol{\psi} = 0$. This is, of course, the Euler-Lagrange equation $\boldsymbol{\psi} = 0$.

J.2 Second Variation

The differential of the second variation of the action $S[\mathbf{q}]$ is

$$\begin{aligned} \delta^2 S[\mathbf{q}, \boldsymbol{\eta}] &= \int_{t_a}^{t_b} \left(\boldsymbol{\eta}^T \cdot \frac{\partial^2 L}{\partial \mathbf{q} \partial \mathbf{q}} \cdot \boldsymbol{\eta} + 2 \boldsymbol{\eta}^T \cdot \frac{\partial^2 L}{\partial \mathbf{q} \partial \dot{\mathbf{q}}} \cdot \dot{\boldsymbol{\eta}} + \right. \\ &\quad \left. \dot{\boldsymbol{\eta}}^T \cdot \frac{\partial^2 L}{\partial \dot{\mathbf{q}} \partial \dot{\mathbf{q}}} \cdot \dot{\boldsymbol{\eta}} \right) (t) dt. \end{aligned} \quad (137)$$

For $t \mapsto \mathbf{q}(t)$ not only an extremum of $S[\mathbf{q}]$, where we already know that $\delta S[\mathbf{q}, \boldsymbol{\eta}] = 0$ for all $\boldsymbol{\eta}$, but a *minimum*, we have the condition that, for all variations $t \mapsto \boldsymbol{\eta}(t)$,

$$\delta^2 S[\mathbf{q}, \boldsymbol{\eta}] \geq 0. \quad (138)$$

Using integration by parts, we can obtain the Legendre form

$$\delta^2 S[\mathbf{q}, \boldsymbol{\eta}] = \int_{t_a}^{t_b} \left(\dot{\boldsymbol{\eta}}^T \cdot \mathbf{P} \cdot \dot{\boldsymbol{\eta}} + \boldsymbol{\eta}^T \cdot \mathbf{Q} \cdot \boldsymbol{\eta} \right) (t) dt, \quad (139)$$

with the terms

$$\begin{aligned} \mathbf{P}(t) &= \frac{\partial^2 L}{\partial \dot{\mathbf{q}} \partial \dot{\mathbf{q}}} \Big|_{(\mathbf{q}(t), \dot{\mathbf{q}}(t), t)}, \\ \mathbf{Q}(t) &= \left[\frac{\partial^2 L}{\partial \mathbf{q} \partial \mathbf{q}} + \frac{d}{dt} \frac{\partial^2 L}{\partial \mathbf{q} \partial \dot{\mathbf{q}}} \right]_{(\mathbf{q}(t), \dot{\mathbf{q}}(t), t)}. \end{aligned} \quad (140)$$

K FREE BOUNDARIES AND NATURAL BOUNDARY CONDITIONS

A free boundary corresponding to a fixed time $t = \tau$ (e.g., $\tau = t_a$ or $\tau = t_b$, for $t \in [t_a, t_b]$) results from the *natural boundary conditions*

$$\frac{\partial L}{\partial \dot{\mathbf{q}}} = \mathbf{p} = 0. \quad (141)$$

These natural boundary conditions come from the derivation of the functional differential in Eq. 130, which shows that if they are fulfilled the value of the functional differential $\delta S[\mathbf{q}, \boldsymbol{\eta}]$ will always vanish, even for non-vanishing variation $\boldsymbol{\eta}(\tau)$ at the boundary $t = \tau$.

K.1 The Subspace of a Free Boundary

Corresponding to a natural boundary condition at boundary $t = \tau$, all possible optimal solutions $(\mathbf{q}, \dot{\mathbf{q}})$ span a k -dimensional linear subspace of the $2k$ -dimensional phase space $(\mathbf{q}, \dot{\mathbf{q}}, \tau)$. From Eq. 25a, we obtain that an optimal solution with $\mathbf{p} = 0$ must fulfill Hamilton's equation

$$\dot{\mathbf{q}}(\mathbf{q}, \tau) = -L_{\dot{\mathbf{q}}}^{-1}(\tau) (L_{\mathbf{q}\mathbf{q}}(\tau) \mathbf{q} - \bar{\mathbf{b}}_l(\tau)). \quad (142)$$

We therefore can simply parameterize the k -dimensional subspace of optimal solutions via all tuples $\mathbf{q} = (q^1, \dots, q^k)^T$. Then, the solution subspace, for time $t = \tau$, is explicitly given by all affine combinations

$$\begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + q^1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + q^k \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mathbf{b}_k \end{pmatrix}. \quad (143)$$

Each k -vector $\mathbf{b}_i$ is defined by the i 'th column of the $k \times k$ matrix

$$\mathbf{B}(\tau) = -L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}^{-1} L_{\mathbf{q}\mathbf{q}}(\tau). \quad (144)$$

Per Eq. 142, the affine part of Eq. 143 contains the k -vector

$$\dot{\mathbf{q}}(0, \tau) = L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}^{-1} \bar{\mathbf{b}}_l(\tau). \quad (145)$$

K.2 Functional Gradient Descent with Free Boundary

From Eqs. 130 and 141, we can derive that a free boundary at times t_a and t_b , respectively, can be computed via functional gradient descent with the functional gradient $\boldsymbol{\psi}(t)$ at the boundary given by

$$\begin{aligned} \boldsymbol{\psi}(t_a) &= -\mathbf{p}(t_a), \\ \boldsymbol{\psi}(t_b) &= \mathbf{p}(t_b). \end{aligned} \quad (146)$$

We note one important caveat: When the initial curve $t \mapsto \mathbf{q}(t)$ from which functional gradient descent is started is chosen via point-wise computation using Eq. 34, it is crucial to realize that $\mathbf{p}(t_a)$ and $\mathbf{p}(t_b)$, respectively, will in general *not be zero*, although one might mistakenly think that this is the case. Their actual values must be computed from the initial curve via Eq. 36 and the definition $\mathbf{p} = \partial L / \partial \dot{\mathbf{q}}$. That is,

$$\begin{aligned} \dot{\mathbf{q}}(t_a) &= \frac{d}{dt} \Big|_{t=t_a} \mathbf{q}(t), \quad \mathbf{p}(t_a) = \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}(t_a), \dot{\mathbf{q}}(t_a), t_a), \\ \dot{\mathbf{q}}(t_b) &= \frac{d}{dt} \Big|_{t=t_b} \mathbf{q}(t), \quad \mathbf{p}(t_b) = \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}(t_b), \dot{\mathbf{q}}(t_b), t_b). \end{aligned} \quad (147)$$

K.3 Functional Gradient Descent with Fixed Boundary

For a fixed boundary, where $\mathbf{q}(t_a)$ and $\mathbf{q}(t_b)$ are fixed and are not allowed to change, the functional gradient at the boundary is simply

$$\begin{aligned} \boldsymbol{\psi}(t_a) &= 0, \\ \boldsymbol{\psi}(t_b) &= 0. \end{aligned} \quad (148)$$

Here, it is again important to note that the caveat above, together with the mandatory requirement of fulfilling Eq. 147, equally applies, even though $\mathbf{p}(t_a), \mathbf{p}(t_b)$ are not used in the functional gradient at the boundary. The reason for this is that $\boldsymbol{\psi}(t)$ with $t \in (t_a, t_b)$, i.e., in the interior of the curve, must still be computed with the correct derivatives $\dot{\mathbf{p}}(t)$, which depend on $\mathbf{p}(t_a), \mathbf{p}(t_b)$ at the boundary.

L EXPLICIT COMPUTATION OF LAGRANGIANS

In this appendix, we give explicit computations for several variants of using our framework for Lagrangians L of the form given by Eq. 11:

$$L(\mathbf{q}, \dot{\mathbf{q}}, t) = \frac{1}{2} \left\| \mathbf{A}(t)(\mathbf{q}, \dot{\mathbf{q}})^T - \mathbf{b}(t) \right\|^2.$$

For discretization, we define a *grid resolution* with a corresponding number N of grid points x . Moreover, for approximate Killing fields $\mathbf{u}$, we define the number of *degrees of freedom* of the field $\mathbf{u}$ per time t as $k = n \cdot \tilde{N}$, with a *number of nodes* $\tilde{N}$. Here, $\tilde{N}$ is separated from N to allow for fewer nodes, and thus fewer degrees of freedom, for the approximate Killing field $\mathbf{u}$ than there are actual grid points. By default, we choose $\tilde{N} = N$. However, we can also choose $\tilde{N} < N$, and incorporate spatial basis functions that re-construct the approximate Killing field values at all N grid points from just $\tilde{N}$ node values.

L.1 Observed Time Derivative Lagrangian

We derive explicit equations for the Lagrangian L_x^D given by Eq. 10a for the case where, unlike for the path line-constrained approach defined in Appendix L.2, we do not place any additional restrictions on the number of degrees of freedom for the observer field. For a Euclidean observer, we thus have k degrees of freedom with $k = 3$ for 2D flows, $k = 6$ for 3D flows. In this case, we can always use a time-independent basis $\mathbf{e}_i$, and we therefore can simplify Eq. 10a for the case $\partial \mathbf{e}_i / \partial t = 0$:

$$\begin{aligned} L_x^D(\mathbf{q}, \dot{\mathbf{q}}, t) &= \left\| q^i \left(\nabla \mathbf{v}(\mathbf{e}_i) - \nabla \mathbf{e}_i(\mathbf{v}) - \frac{\partial \mathbf{e}_i}{\partial t} \right) - \dot{q}^i \mathbf{e}_i - \mathbf{b} \right\|^2(x, t), \\ &= \left\| q^i \left(\nabla \mathbf{v}(\mathbf{e}_i) - \nabla \mathbf{e}_i(\mathbf{v}) \right) - \dot{q}^i \mathbf{e}_i + \frac{\partial \mathbf{v}}{\partial t} \right\|^2(x, t). \end{aligned} \quad (149)$$

Moreover, in this case, the observed time derivative is given by simply choosing $\mathbf{b}(x, t) = -\partial \mathbf{v}(x, t) / \partial t$. We now first derive explicit equations for Euclidean observers modeled as Killing fields, and then derive the equations for observers modeled as approximate Killing fields.

L.1.1 Killing fields

In order to generate Killing vector fields $\mathbf{u}$ from k coefficients q^i , we define the following matrices $\mathbf{E}$ containing all necessary basis Killing fields $\mathbf{e}_i$ in their columns, and matrices mapping the coefficients q^i to the corresponding term $\nabla \mathbf{e}_i(\mathbf{v})$, with $\mathbf{v} = \sum_{i=1}^n v^i \mathbf{e}_i$, respectively:

$$2D: \mathbf{E}(x) = \begin{pmatrix} 1 & 0 & -x^2 \\ 0 & 1 & x^1 \end{pmatrix} (x), \quad (150)$$

$$\nabla \mathbf{E} \mathbf{V}(x, t) = \begin{pmatrix} 0 & 0 & -v^2 \\ 0 & 0 & v^1 \end{pmatrix} (x, t), \quad (151)$$

$$3D: \mathbf{E}(x) = \begin{pmatrix} 1 & 0 & 0 & 0 & x^3 & -x^2 \\ 0 & 1 & 0 & -x^3 & 0 & x^1 \\ 0 & 0 & 1 & x^2 & -x^1 & 0 \end{pmatrix} (x), \quad (152)$$

$$\nabla \mathbf{E} \mathbf{V}(x, t) = \begin{pmatrix} 0 & 0 & 0 & 0 & v^3 & -v^2 \\ 0 & 0 & 0 & -v^3 & 0 & v^1 \\ 0 & 0 & 0 & v^2 & -v^1 & 0 \end{pmatrix} (x, t). \quad (153)$$

We obtain $\mathbf{u}(x, t) = \mathbf{E}(x) \mathbf{q}(t)$, and $\nabla \mathbf{u}(\mathbf{v}) = \nabla \mathbf{E} \mathbf{V}(x, t) \mathbf{q}(t)$.

We then define the $k \times k$ matrices

$$\mathbf{G} = \int_U \left(\mathbf{E}^T \mathbf{E} \right) (x) dx. \quad (154)$$

We do not use this fact explicitly, but note that $\mathbf{G}$ contains the moment of inertia tensor around the center of mass of the domain U for equal mass distribution, and that $\mathbf{q}^T \mathbf{G} \mathbf{q} = \int_U (\mathbf{u}^T \mathbf{u})(x) dx$.

We furthermore define matrices $\mathbf{L}(x, t)$ for time-independent Lie derivatives $\mathcal{L}_{\mathbf{u}} \mathbf{v}$, i.e., Lie brackets $[\mathbf{u}, \mathbf{v}]$. We define the $n \times k$ matrices

$$\mathbf{L}(x, t) = \nabla \mathbf{v}(x, t) \mathbf{E}(x) - \nabla \mathbf{E} \mathbf{V}(x, t). \quad (155)$$

At (x, t) , we have $\mathcal{L}_{\mathbf{u}} \mathbf{v} = [\mathbf{u}, \mathbf{v}] = \nabla \mathbf{v}(\mathbf{u}) - \nabla \mathbf{u}(\mathbf{v}) = \mathbf{L}(x, t) \mathbf{q}(t)$.

Furthermore, we define the $n \times 2k$ matrices

$$\mathbf{C}(x, t) = \left(\mathbf{L}(x, t) \mid -\mathbf{E}(x) \right). \quad (156)$$

We can now compute observed time derivative n -vectors as

$$\begin{aligned} \mathbf{d}(x, t) &= \mathbf{L}(x, t) \mathbf{q} - \mathbf{E}(x) \dot{\mathbf{q}} + \frac{\partial \mathbf{v}}{\partial t}(x, t), \\ &= \mathbf{C}(x, t) \begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix} + \frac{\partial \mathbf{v}}{\partial t}(x, t). \end{aligned} \quad (157)$$

The integrated Lagrangian $L(\mathbf{q}, \dot{\mathbf{q}}, t)$ then is

$$L(\mathbf{q}, \dot{\mathbf{q}}, t) = \frac{1}{2} \int_U \mathbf{d}(x, t)^T \mathbf{d}(x, t) dx. \quad (158)$$

Considering the general form of the Lagrangian L in Eq. 11, the $m \times 2k$ matrix $\mathbf{A}(t)$ is determined by vertically stacking the matrices $\mathbf{C}(x, t)$ for all grid points x . Likewise, the m -vector $\mathbf{b}(t)$ is determined by vertically stacking all partial time derivatives of the input field $\mathbf{v}$. If the number of grid points is N , the height m of the matrix $\mathbf{A}(t)$, and the size of the vector $\mathbf{b}(t)$, respectively, is $m = n \cdot N$. However, instead of first computing $\mathbf{A}(t)$ and then computing $\mathbf{A}^T \mathbf{A}(t)$, it is simpler to directly compute the $2k \times 2k$ Hessian matrix for each time t as

$$\mathbf{H}_L(t) = \begin{pmatrix} L_{\mathbf{q}\mathbf{q}}(t) & L_{\mathbf{q}\dot{\mathbf{q}}}(t) \\ L_{\dot{\mathbf{q}}\mathbf{q}}(t) & L_{\dot{\mathbf{q}}\dot{\mathbf{q}}}(t) \end{pmatrix} = \int_U \left(\mathbf{C}^T \mathbf{C} \right) (x, t) dx. \quad (159)$$

The four $k \times k$ Hessian matrix sub-blocks above are therefore

$$\begin{aligned} L_{\mathbf{q}\mathbf{q}}(t) &= \int_U \left(\mathbf{L}^T \mathbf{L} \right) (x, t) dx, \\ L_{\mathbf{q}\dot{\mathbf{q}}}(t) &= - \int_U \left(\mathbf{L}^T \mathbf{E} \right) (x, t) dx, \\ L_{\dot{\mathbf{q}}\mathbf{q}}(t) &= - \int_U \left(\mathbf{E}^T \mathbf{L} \right) (x, t) dx, \\ L_{\dot{\mathbf{q}}\dot{\mathbf{q}}} &= \int_U \left(\mathbf{E}^T \mathbf{E} \right) (x) dx = \mathbf{G}. \end{aligned} \quad (160)$$

The Hessian matrix in the last row above, i.e., $\mathbf{G}$, does not depend on the vector field $\mathbf{v}$ at all. If we assume that here the domain U is not time-dependent (unlike the path line-constrained approach in Appendix L.2), the Hessian $\mathbf{G}$ is furthermore even time-independent.

The $2k$ -vector right-hand side $\mathbf{A}^T \mathbf{b}(t)$ can be computed directly as

$$\mathbf{A}^T \mathbf{b}(t) = - \int_U \mathbf{C}^T \frac{\partial \mathbf{v}}{\partial t}(x, t) dx. \quad (161)$$

The upper k -vector right-hand side $\bar{\mathbf{b}}_h(t) = \mathbf{P}_h \mathbf{A}^T \mathbf{b}(t)$ thus is

$$\bar{\mathbf{b}}_h(t) = - \int_U \mathbf{L}^T \frac{\partial \mathbf{v}}{\partial t}(x, t) dx. \quad (162)$$

The lower k -vector right-hand side $\bar{\mathbf{b}}_l(t) = \mathbf{P}_l \mathbf{A}^T \mathbf{b}(t)$ thus is

$$\bar{\mathbf{b}}_l(t) = \int_U \mathbf{E}^T \frac{\partial \mathbf{v}}{\partial t}(x, t) dx. \quad (163)$$

With the above, we can compute the partial derivatives

$$L_{\mathbf{q}}(\mathbf{q}, \dot{\mathbf{q}}, t) = L_{\mathbf{q}\mathbf{q}}(t) \mathbf{q} - L_{\mathbf{q}\dot{\mathbf{q}}}(t) \dot{\mathbf{q}} - \bar{\mathbf{b}}_h(t). \quad (164)$$

Second-order time derivatives

A big strength of our Lagrangian-Hamiltonian method is that *no second-order time derivatives are required* to evaluate the Euler-Lagrange equation. In contrast, for the purely Lagrangian approach, because of Eq. 66, we would need to compute the second-order partial derivatives

$$\frac{\partial^2 L}{\partial t \partial \dot{\mathbf{q}}} = - \frac{\partial}{\partial t} \int_U \left(\mathbf{E}^T \mathbf{d} \right) (x, t) dx. \quad (165)$$

Because $\mathbf{E}$ is time-independent, we can expand this as

$$\begin{aligned} \frac{\partial^2 L}{\partial t \partial \dot{\mathbf{q}}} &= - \int_U \mathbf{E}^T(x) \left(\left(\frac{\partial}{\partial t} \mathbf{L}(x, t) \right) \mathbf{q} + \frac{\partial^2 \mathbf{v}}{\partial t^2}(x, t) \right) dx, \\ &= - \left(\int_U \mathbf{E}^T(x) \left(\frac{\partial}{\partial t} \mathbf{L}(x, t) \right) dx \right) \mathbf{q} \\ &\quad - \int_U \mathbf{E}^T(x) \frac{\partial^2 \mathbf{v}}{\partial t^2}(x, t) dx. \end{aligned} \quad (166)$$

L.1.2 Approximate Killing fields

Considering the general form of the Lagrangian L in Eq. 11, we need to set up the $m \times 2k$ matrix $\mathbf{A}(t)$, and the m -vector right-hand side $\mathbf{b}(t)$, respectively. For an approximate Killing field $\mathbf{u}$, we choose $m = n \cdot \tilde{N}$, where $\tilde{N}$ is the number of grid points, and $n = 2$ for 2D flows, and $n = 3$ for 3D flows, respectively. Moreover, $k = n \cdot \tilde{N}$ is the number of degrees of freedom of the field $\mathbf{u}$, where we either choose $\tilde{N} = N$, or $\tilde{N} < N$. To obtain $\mathbf{A}(t)$, $\mathbf{b}(t)$, we vertically stack $\tilde{N}$ of the $n \times 2k$ matrices $\mathbf{A}(x, t)$, and the n -vectors $\mathbf{b}(x, t)$, respectively, given by

$$\begin{aligned} \mathbf{A}(x, t) &= \left(\left(\nabla \mathbf{v}(x, t) \cdot \mathbf{E}(x) \right) - \left(\nabla \mathbf{E} \mathbf{V} \right)(x, t) \mid -\mathbf{E}(x) \right), \\ \mathbf{b}(x, t) &= - \frac{\partial \mathbf{v}}{\partial t}(x, t). \end{aligned} \quad (167)$$

If $\tilde{N} = N$, i.e., we use as many nodes as there are grid points x , for simplicity we can use spatial box basis functions, by defining the $n \times k$ matrices $\mathbf{E}(x)$ with an identity matrix in position x , i.e.,

$$\mathbf{E}(x) = \left(\dots \mathbf{0} \dots \mid \mathbf{I}^{n \times n} \mid \dots \mathbf{0} \dots \right). \quad (168)$$

If $\tilde{N} < N$, instead of identity matrices, we have to incorporate spatial basis functions that map the coefficients of $\tilde{N}$ nodes to N grid points.

The tensor field $\nabla \mathbf{u}$ can be computed using the operator $\mathbf{D}_K: \mathbb{R}^k \rightarrow \mathbb{R}^{N \cdot n \cdot n}$ from Eq. 211. The $n \times k$ matrices $\nabla \mathbf{E} \mathbf{V}$ are

$$\left(\nabla \mathbf{E} \mathbf{V} \right)(x, t) = \mathbf{F}_2 \cdot \mathbf{F}_1(\mathbf{v}, x, t) \cdot \mathbf{D}_K. \quad (169)$$

The operator $\mathbf{F}_1(\mathbf{v}, x, t): \mathbb{R}^{N \cdot n \cdot n} \rightarrow \mathbb{R}^{n \cdot n}$, as a $n^2 \times Nn^2$ matrix, contains the replicated n -vector components $v^j(x, t)$ at grid point x , in an $n^2 \times n^2$ diagonal matrix block such that by right-multiplication with $\mathbf{D}_K$, and left-multiplication by $\mathbf{F}_2: \mathbb{R}^{n \cdot n} \rightarrow \mathbb{R}^n$, given by a $n \times n^2$ matrix, the n -vector $\nabla \mathbf{u}(\mathbf{v})$, at point x , will result.

For example, in 2D ($n = 2$), assuming that $\mathbf{D}_K$ produces elements in row-major order (cf. Eq. 212), we have the $n^2 \times Nn^2$ matrix

$$\begin{aligned} \mathbf{F}_1(\mathbf{v}, x, t) &= \\ &\left(\dots, \mathbf{0}, \text{diag}(v^1, v^2, v^1, v^2)(x, t), \mathbf{0}, \dots \right). \end{aligned} \quad (170)$$

Multiplied by $\mathbf{D}_K \cdot \mathbf{q}$ on the right, at x we thus obtain the n^2 -vector

$$\mathbf{F}_1(\mathbf{v}, x, t) \cdot \mathbf{D}_K \cdot \mathbf{q} = \begin{pmatrix} v^1 \cdot \nabla_1 u^1 \\ v^2 \cdot \nabla_2 u^1 \\ v^1 \cdot \nabla_1 u^2 \\ v^2 \cdot \nabla_2 u^2 \end{pmatrix} (x, t). \quad (171)$$

The operator $\mathbf{F}_2: \mathbb{R}^{n \cdot n} \rightarrow \mathbb{R}^n$ then maps this to the final n -vector for point x by adding together runs of n consecutive rows each, giving

$$\mathbf{F}_2 \cdot \mathbf{F}_1(\mathbf{v}, x, t) \cdot \mathbf{D}_K \cdot \mathbf{q} = \left(\sum_{i=1}^n v^i \cdot \nabla_i u^1, \sum_{i=1}^n v^i \cdot \nabla_i u^2 \right) (x, t). \quad (172)$$

L.2 Euclidean Path Line Observer Lagrangian

By defining an appropriate Lagrangian, our framework can immediately be used to determine optimal Euclidean observers that are constrained to one specific (but arbitrarily chosen) path line $t \mapsto c(t) \in \mathbb{R}^n$ of the field $\mathbf{v}(x, t)$, as introduced by Günther and Theisel [14]. (We denote the path line by $c(t)$ to avoid confusion with x or p as they are already used.) Constraining the observer removes n degrees of freedom from the observer Killing field, and we therefore have $k = 1$ for 2D input flow fields, and $k = 3$ for 3D input flow fields. Moreover, this constraint also implies that here, unlike for the time derivative defined in Appendix L.1, we *cannot* use a time-independent basis $\mathbf{e}_i$. We will therefore have $\partial \mathbf{e}_i / \partial t \neq 0$. For this reason, the Lagrangian L_x^D in Eq. 10a must be computed using the fully general expression

$$L_x^D(\mathbf{q}, \dot{\mathbf{q}}, t) = \left\| \dot{q}^i \left(\nabla \mathbf{v}(\mathbf{e}_i) - \nabla \mathbf{e}_i(\mathbf{v}) - \frac{\partial \mathbf{e}_i}{\partial t} \right) - \dot{q}^i \mathbf{e}_i - \mathbf{b} \right\|^2(x, t).$$

See Eq. 182 for the term $\mathbf{b}(x, t)$. For 2D and 3D flows, respectively, due to the path line constraint the generalized coordinates $\mathbf{q}$ now are¹²

$$\text{2D: } \mathbf{q} = q^3, \quad \text{3D: } \mathbf{q} = (q^4, q^5, q^6)^T. \quad (173)$$

Here, the coordinate $\mathbf{q}$ only encodes the angular velocity of the observer field, with the generalized velocity $\dot{\mathbf{q}}$ its rate of change. Given a specific path line $t \mapsto c(t)$, we define the following $n \times k$ matrices. We denote the coordinates of points $x \in \mathbb{R}^n$ as $x^i = x^i(x)$, and those of points on the path line as $c^i(t) = x^i(c(t))$. For the input field $\mathbf{v} = \sum_{i=1}^n v^i \mathbf{e}_i$,

$$\text{2D: } \mathbf{E}(x, t) = \begin{pmatrix} -(x^2 - c^2) \\ x^1 - c^1 \end{pmatrix} (x, t), \quad (174)$$

$$\mathbf{E}_t(t) = \begin{pmatrix} v^2 \\ -v^1 \end{pmatrix} (c(t), t), \quad (175)$$

$$\nabla \mathbf{E} \mathbf{V}(x, t) = \begin{pmatrix} -v^2 \\ v^1 \end{pmatrix} (x, t), \quad (176)$$

$$\text{3D: } \mathbf{E}(x) = \begin{pmatrix} 0 & x^3 - c^3 & -(x^2 - c^2) \\ -(x^3 - c^3) & 0 & x^1 - c^1 \\ x^2 - c^2 & -(x^1 - c^1) & 0 \end{pmatrix} (x, t), \quad (177)$$

$$\mathbf{E}_t(t) = \begin{pmatrix} 0 & -v^3 & v^2 \\ v^3 & 0 & -v^1 \\ -v^2 & v^1 & 0 \end{pmatrix} (c(t), t), \quad (178)$$

$$\nabla \mathbf{E} \mathbf{V}(x, t) = \begin{pmatrix} 0 & v^3 & -v^2 \\ -v^3 & 0 & v^1 \\ v^2 & -v^1 & 0 \end{pmatrix} (x, t). \quad (179)$$

As in Appendix L.1.1, here we have $\nabla \mathbf{u}(\mathbf{v}) = \nabla \mathbf{E} \mathbf{V}(x, t) \mathbf{q}$, but due to fewer degrees of freedom the Killing field $\mathbf{u}$ is now given by

$$\mathbf{u}(x, t) = \mathbf{E}(x, t) \mathbf{q} + \mathbf{v}(c(t), t). \quad (180)$$

Thus, here the Eulerian time derivative of the Killing field $\mathbf{u}$ is

$$\frac{\partial \mathbf{u}}{\partial t}(x, t) = \mathbf{E}(x, t) \dot{\mathbf{q}} + \mathbf{E}_t(t) \mathbf{q} + \left(\nabla \mathbf{v}(\mathbf{v}) + \frac{\partial \mathbf{v}}{\partial t} \right) (c(t), t). \quad (181)$$

¹²To avoid confusion between this and our other Lagrangians, here we denote the degrees of freedom such that they correspond to angular velocity in our other Lagrangians. That is, in 2D by q^3 instead of by q^1 , and in 3D by (q^4, q^5, q^6) .

Thus, the observed time derivative at (x, t) is given by the n -vector

$$\begin{aligned} \mathbf{d}(x, t) &= \mathbf{A}(x, t) \begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix} - \mathbf{b}(x, t), \\ \mathbf{A}(x, t) &= \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \end{pmatrix} (x, t), \\ \mathbf{A}_1(x, t) &= \nabla \mathbf{v}(x, t) \mathbf{E}(x, t) - \nabla \mathbf{E} \mathbf{V}(x, t) - \mathbf{E}_t(t), \\ \mathbf{A}_2(x, t) &= -\mathbf{E}(x, t), \\ \mathbf{b}(x, t) &= (\nabla \mathbf{v}(c(t), t) - \nabla \mathbf{v}(x, t)) \mathbf{v}(c(t), t) \\ &\quad + \frac{\partial \mathbf{v}}{\partial t}(c(t), t) - \frac{\partial \mathbf{v}}{\partial t}(x, t). \end{aligned} \quad (182)$$

Lagrangian. With Eq. 182, the integrated Lagrangian L is

$$L(\mathbf{q}, \dot{\mathbf{q}}, t) = \frac{1}{2} \int_{U(t)} \mathbf{d}(x, t)^T \mathbf{d}(x, t) dx. \quad (183)$$

L.2.1 Full example computation: 2D flow fields

For unsteady 2D flow fields ($n = 2$), the 2×2 matrix $\mathbf{A}(x, t)$ consists only of two 2D column vectors $\mathbf{a}_1(x, t)$, $\mathbf{a}_2(x, t)$, i.e.,

$$\begin{aligned} \mathbf{A}_1(x, t) &= \mathbf{a}_1(x, t), \\ \mathbf{A}_2(x, t) &= \mathbf{a}_2(x, t). \end{aligned} \quad (184)$$

We will now abbreviate $q = q^3$, $\dot{q} = \dot{q}^3$, $p = p_3$. Thus,

$$\mathbf{d}(x, t) = \mathbf{a}_1(x, t) q + \mathbf{a}_2(x, t) \dot{q} - \mathbf{b}(x, t). \quad (185)$$

If we stack $\mathbf{a}_1(x, t)$, $\mathbf{a}_2(x, t)$, respectively, for all x , we obtain the columns $\mathbf{a}_1(t)$, $\mathbf{a}_2(t)$ of the matrix $\mathbf{A}(t)$ in Eq. 11. Likewise, for the right-hand side, stacking $\mathbf{b}(x, t)$ for all x gives $\mathbf{b}(t)$ in Eq. 11. Thus,

$$\begin{aligned} \mathbf{a}_1^T \mathbf{a}_1(t) &= \int_{U(t)} \mathbf{a}_1^T \mathbf{a}_1(x, t) dx, \\ \mathbf{a}_1^T \mathbf{a}_2(t) &= \int_{U(t)} \mathbf{a}_1^T \mathbf{a}_2(x, t) dx, \\ \mathbf{a}_2^T \mathbf{a}_2(t) &= \int_{U(t)} \mathbf{a}_2^T \mathbf{a}_2(x, t) dx, \\ \mathbf{a}_1^T \mathbf{b}(t) &= \int_{U(t)} \mathbf{a}_1^T \mathbf{b}(x, t) dx, \\ \mathbf{a}_2^T \mathbf{b}(t) &= \int_{U(t)} \mathbf{a}_2^T \mathbf{b}(x, t) dx, \\ \mathbf{b}^T \mathbf{b}(t) &= \int_{U(t)} \mathbf{b}^T \mathbf{b}(x, t) dx. \end{aligned} \quad (186)$$

For the integrated Lagrangian L from Eq. 183 we thus get

$$\begin{aligned} L(q, \dot{q}, t) &= \frac{1}{2} \mathbf{a}_1^T \mathbf{a}_1(t) (q)^2 + \frac{1}{2} \mathbf{a}_2^T \mathbf{a}_2(t) (\dot{q})^2 + \frac{1}{2} \mathbf{b}^T \mathbf{b}(t) \\ &\quad + \mathbf{a}_1^T \mathbf{a}_2(t) q \dot{q} - \mathbf{a}_1^T \mathbf{b}(t) q - \mathbf{a}_2^T \mathbf{b}(t) \dot{q}, \\ L_q &= q \mathbf{a}_1^T \mathbf{a}_1(t) + \dot{q} \mathbf{a}_1^T \mathbf{a}_2(t) - \mathbf{a}_1^T \mathbf{b}(t), \\ p = L_{\dot{q}} &= q \mathbf{a}_1^T \mathbf{a}_2(t) + \dot{q} \mathbf{a}_2^T \mathbf{a}_2(t) - \mathbf{a}_2^T \mathbf{b}(t), \\ L_{qq} &= \mathbf{a}_1^T \mathbf{a}_1(t), \\ L_{q\dot{q}} &= \mathbf{a}_1^T \mathbf{a}_2(t), \\ L_{\dot{q}\dot{q}} &= \mathbf{a}_2^T \mathbf{a}_2(t). \end{aligned} \quad (187)$$

As in Eq. 60, we can also write the Lagrangian L as

$$L(q, \dot{q}, t) = \frac{1}{2} \begin{pmatrix} q \\ \dot{q} \end{pmatrix}^T \mathbf{H}_L(t) \begin{pmatrix} q \\ \dot{q} \end{pmatrix} - \begin{pmatrix} q \\ \dot{q} \end{pmatrix}^T \begin{pmatrix} \mathbf{a}_1^T \mathbf{b} \\ \mathbf{a}_2^T \mathbf{b} \end{pmatrix} (t) + \frac{1}{2} \mathbf{b}^T \mathbf{b}(t). \quad (188)$$

With $p = L_{\dot{q}}$, the Hamiltonian vector field therefore is

$$\begin{aligned}\dot{q}(q, p, t) &= L_{\dot{q}}^{-1} (p - L_{\dot{q}q} q + \bar{b}_l(t)), \\ &= \frac{1}{\mathbf{a}_2^T \mathbf{a}_2(t)} (p - \mathbf{a}_1^T \mathbf{a}_2(t) q + \mathbf{a}_2^T \mathbf{b}(t)), \\ \dot{p}(q, p, t) &= L_q, \\ &= \mathbf{a}_1^T \mathbf{a}_1(t) q + \mathbf{a}_1^T \mathbf{a}_2(t) \dot{q} - \mathbf{a}_1^T \mathbf{b}(t).\end{aligned}\quad (189)$$

The Hamiltonian matrix $\mathbf{X}(t)$ for Eq. 26 is given by

$$\mathbf{X}(t) = \begin{pmatrix} 0 & 0 \\ \mathbf{a}_1^T \mathbf{a}_1 & 0 \end{pmatrix} (t) + \frac{1}{\mathbf{a}_2^T \mathbf{a}_2} \begin{pmatrix} -\mathbf{a}_1^T \mathbf{a}_2 & 1 \\ -(\mathbf{a}_1^T \mathbf{a}_2)^2 & \mathbf{a}_1^T \mathbf{a}_2 \end{pmatrix} (t). \quad (190)$$

The matrix $\mathbf{X}(t)$ has real eigenvalues

$$\lambda = \pm \sqrt{\frac{\mathbf{a}_1^T \mathbf{a}_1}{\mathbf{a}_2^T \mathbf{a}_2}}. \quad (191)$$

The affine part for Eq. 26 is

$$X_H(0, 0, t) = \begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} (0, 0, t) = \frac{1}{\mathbf{a}_2^T \mathbf{a}_2} \begin{pmatrix} \mathbf{a}_2^T \mathbf{b} \\ 0 \end{pmatrix} (t). \quad (192)$$

From Eq. 31, the Hessian matrix of the Hamiltonian H is therefore

$$\begin{aligned}\mathbf{H}_H(t) &= \\ &= \begin{pmatrix} -\mathbf{a}_1^T \mathbf{a}_1 & 0 \\ 0 & 0 \end{pmatrix} (t) + \frac{1}{\mathbf{a}_2^T \mathbf{a}_2} \begin{pmatrix} (\mathbf{a}_1^T \mathbf{a}_2)^2 & -\mathbf{a}_1^T \mathbf{a}_2 \\ -\mathbf{a}_1^T \mathbf{a}_2 & 1 \end{pmatrix} (t).\end{aligned}\quad (193)$$

From Eq. 32, the Hamiltonian H is therefore given by

$$H(q, p, t) = \frac{1}{2} \begin{pmatrix} q \\ p \end{pmatrix}^T \mathbf{H}_H(t) \begin{pmatrix} q \\ p \end{pmatrix} + p \dot{q}(0, 0, t) + \frac{1}{2} \mathbf{b}^T \mathbf{b}(t). \quad (194)$$

It is also interesting to consider the natural boundary conditions from Appendix K explicitly, where the one-dimensional (because $k = 1$) subspace of $(q, \dot{q})$ coordinates allowed at a free boundary at arbitrary time $t = \tau$, is given by Eq. 143. Obtaining the one-dimensional basis vector $\mathbf{b}_1$, and the 1D affine point, respectively, of this subspace via Eqs. 144, 145, we can write this 1D subspace via Eq. 143 as

$$\begin{pmatrix} q \\ \dot{q} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{\mathbf{a}_2^T \mathbf{a}_2} \mathbf{a}_2^T \mathbf{b} \end{pmatrix} (\tau) + q \begin{pmatrix} 1 \\ -\frac{1}{\mathbf{a}_2^T \mathbf{a}_2} \mathbf{a}_1^T \mathbf{a}_2 \end{pmatrix} (\tau). \quad (195)$$

If we explicitly write the general Hamiltonian H from Eq. 32 only for this subspace, i.e., we specialize for the subspace $p = 0$, we have

$$H(q, 0, t) = \frac{1}{2} \left(\frac{1}{\mathbf{a}_2^T \mathbf{a}_2} (\mathbf{a}_1^T \mathbf{a}_2)^2 - \mathbf{a}_1^T \mathbf{a}_1 \right) (t) q^2 + \frac{1}{2} \mathbf{b}^T \mathbf{b}(t). \quad (196)$$

L.3 Matching a Seed Observer Field

We derive explicit equations for the Lagrangian L_x^S in Eq. 10c:

$$L_x^S(\mathbf{q}, \dot{\mathbf{q}}, t) = \left\| \mathbf{s} - q^i \mathbf{e}_i \right\|^2 (x, t).$$

The seed vector field $\mathbf{s}(x, t)$ is a velocity field, either because it is set to $\mathbf{s} := \mathbf{v}$, or because it is set to a known (already computed) observer field. We can minimize the observed field $\mathbf{s} - q^i \mathbf{e}_i$, which directly also minimizes the observed kinetic energy $m \left\| \mathbf{s} - q^i \mathbf{e}_i \right\|^2$, for mass m .

As above, we first derive explicit equations for observers modeled as (exact) Killing fields, and then derive the equations for observers modeled as approximate Killing fields.

L.3.1 Killing fields

We compute the observed velocity field $\mathbf{r}(x, t) = (\mathbf{s} - q^i \mathbf{e}_i)(x, t)$ by

$$\begin{aligned}\mathbf{r}(x, t) &= \mathbf{s}(x, t) - \mathbf{E}(x) \mathbf{q}, \\ &= \mathbf{s}(x, t) + \mathbf{C}(x) \begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix}.\end{aligned}\quad (197)$$

Here, the $n \times 2k$ matrices $\mathbf{C}$ are simply defined by

$$\mathbf{C}(x) = \begin{pmatrix} -\mathbf{E}(x) & \mathbf{0} \end{pmatrix}. \quad (198)$$

The Lagrangian $L(\mathbf{q}, \dot{\mathbf{q}}, t)$, with $L := L^S$, is therefore given by

$$L(\mathbf{q}, \dot{\mathbf{q}}, t) = \frac{1}{2} \int_U \mathbf{r}(x, t)^T \mathbf{r}(x, t) dM. \quad (199)$$

Considering the general form of the Lagrangian L in Eq. 11, the $m \times 2k$ matrix $\mathbf{A}$ is again determined by vertically stacking the matrices $\mathbf{C}(x)$ for all grid points x , and the m -vector $\mathbf{b}(t)$ is determined by the negative of vertically stacking all input field vectors $\mathbf{s}(x, t)$. If the number of grid points is N , the height m of the matrix $\mathbf{A}(t)$, and the size of the vector $\mathbf{b}(t)$, respectively, is $m = n \cdot N$.

We again compute the $2k \times 2k$ Hessian matrices directly as

$$\mathbf{H}_L = \begin{pmatrix} L_{\mathbf{q}\mathbf{q}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \int_U \begin{pmatrix} \mathbf{C}^T \mathbf{C} \end{pmatrix} (x) dx. \quad (200)$$

The top-left Hessian sub-block is

$$L_{\mathbf{q}\mathbf{q}} = \int_U \begin{pmatrix} \mathbf{E}^T \mathbf{E} \end{pmatrix} (x) dx = \mathbf{G}. \quad (201)$$

The $2k$ -vector right-hand side $\mathbf{A}^T \mathbf{b}(t)$ is given by

$$\mathbf{A}^T \mathbf{b}(t) = - \int_U \mathbf{C}^T \mathbf{s}(x, t) dx. \quad (202)$$

The upper and lower k -vector right-hand sides, respectively, thus are

$$\begin{aligned}\bar{\mathbf{b}}_h(t) &= \int_U \mathbf{E}^T \mathbf{s}(x, t) dx, \\ \bar{\mathbf{b}}_l(t) &= 0.\end{aligned}\quad (203)$$

With the above, we can compute the partial derivatives

$$L_{\mathbf{q}}(\mathbf{q}, \dot{\mathbf{q}}, t) = \mathbf{G} \mathbf{q} - \bar{\mathbf{b}}_h(t). \quad (204)$$

L.3.2 Approximate Killing fields

We use the matrices $\mathbf{U}(x)$ from Eq. 168, which extract the components of the velocity vector $\mathbf{u}(x_i)$ at grid point x_i from the vector $\bar{\mathbf{q}}$ as

$$\mathbf{u}(x_i, \tau) = \mathbf{U}(x_i) \bar{\mathbf{T}}(\tau) \bar{\mathbf{D}} \bar{\mathbf{q}}. \quad (205)$$

In the general form of the Lagrangian L in Eq. 11, the $(n \cdot N) \times 2k$ ($k = n \cdot \tilde{N}$) matrix $\mathbf{A}(t)$ is given by vertically stacking N of the matrices $-\mathbf{U}(x_i)$ for N grid points x_i . The right-hand side $\mathbf{b}(t)$ is given by vertically stacking N vectors $-\mathbf{s}(x_i, t)$ for N grid points x_i .

We again compute the $2k \times 2k$ Hessian matrices directly as

$$\mathbf{H}_L = \sum_i \begin{pmatrix} \mathbf{U}(x_i)^T \mathbf{U}(x_i) \end{pmatrix}. \quad (206)$$

The top-left $k \times k$ block of the Hessian matrix is an identity matrix:

$$L_{\mathbf{q}\mathbf{q}} = \mathbf{I}^{k \times k}. \quad (207)$$

All $k \times k$ Hessian matrices except the top-left block are zero, i.e.,

$$L_{\mathbf{q}\dot{\mathbf{q}}} = L_{\dot{\mathbf{q}}\mathbf{q}} = L_{\dot{\mathbf{q}}\dot{\mathbf{q}}} = 0. \quad (208)$$

If the $\mathbf{U}(x_i)$ contain only identity matrix blocks (i.e., if there are no complicated spatial basis functions involved because $N \neq \tilde{N}$), the $2k$ right-hand side vector $\mathbf{A}^T \mathbf{b}(t)$ is simply given by

$$\begin{aligned}\bar{\mathbf{b}}_h(t) &= \mathbf{s}(x, t), \\ \bar{\mathbf{b}}_l(t) &= 0.\end{aligned}\quad (209)$$

L.4 Killing Energy Lagrangian

This energy is only required for observers modeled as approximate Killing fields, since for (exact) Killing fields it is always known to be exactly zero by construction.¹³

For the Lagrangian L_x^K in Eq. 10b, we consider the Killing energy

$$L_x^K(\mathbf{q}, \dot{\mathbf{q}}, t) = \left\| q^i \left(\nabla \mathbf{e}_i + (\nabla \mathbf{e}_i)^T \right) \right\|_F^2(x, t).$$

In discretized form, we define spatial derivative operators $\mathbf{D}_K, \mathbf{D}_{K^T}$ that approximate the Killing energy integrated over a domain U as

$$\begin{aligned} \frac{1}{2} \int_U \left\| \left(q^i \left(\nabla \mathbf{e}_i + (\nabla \mathbf{e}_i)^T \right) \right) (x, t) \right\|_F^2 dx \\ \approx \|(\mathbf{D}_K + \mathbf{D}_{K^T}) \bar{\mathbf{q}}(t)\|^2. \end{aligned} \quad (210)$$

To compute the Frobenius norm of a second-order tensor as the standard L_2 norm of a vector, both derivative operators map k -vectors, with $k = n \cdot \tilde{N}$, to Nn^2 -vectors, with N the number of grid points. That is,

$$\begin{aligned} \mathbf{D}_K: \mathbb{R}^k &\rightarrow \mathbb{R}^{N \cdot n \cdot n}, \\ \mathbf{D}_{K^T}: \mathbb{R}^k &\rightarrow \mathbb{R}^{N \cdot n \cdot n}. \end{aligned} \quad (211)$$

For $\tilde{N} = N$ with spatial box basis functions, these operators can be constructed by using simple finite difference stencils in space. For $\tilde{N} < N$ with general spatial basis functions, these operators must be the derivatives of the basis functions. In both cases, the above operators construct all n^2 derivatives comprising each $\nabla \mathbf{e}_i, (\nabla \mathbf{e}_i)^T$, respectively, component-wise, one row per component, where $\mathbf{D}_K$ and $\mathbf{D}_{K^T}$ only differ in that some rows (those not corresponding to the diagonal of $\nabla \mathbf{e}_i$) are switched, corresponding to the transpose of $\nabla \mathbf{e}_i$.

We can also understand these derivative operator matrices by denoting the observer field by $\mathbf{u} = q^i \mathbf{e}_i$, and considering that $\mathbf{D}_K \bar{\mathbf{q}}$ must construct the tensor field $\nabla \mathbf{u}$. We use row-major order for the sequence of matrix components $\nabla_j u^i$ such that, for example in 2D, we get

$$\mathbf{D}_K \bar{\mathbf{q}} = \begin{pmatrix} \vdots \\ (\nabla_1 u^2)(x_{i-1}, t) \\ (\nabla_2 u^2)(x_{i-1}, t) \\ (\nabla_1 u^1)(x_i, t) \\ (\nabla_2 u^1)(x_i, t) \\ (\nabla_1 u^2)(x_i, t) \\ (\nabla_2 u^2)(x_i, t) \\ (\nabla_1 u^1)(x_{i+1}, t) \\ (\nabla_2 u^1)(x_{i+1}, t) \\ \vdots \end{pmatrix}. \quad (212)$$

To be able to use the general form of the Lagrangian L in Eq. 11, we consider versions of these operators that accept $(\mathbf{q}, \dot{\mathbf{q}})$ arguments instead of $\mathbf{q}$ arguments, but otherwise produce the exact same result. That is, they ignore (i.e., multiply by zero) the $\dot{\mathbf{q}}$ part of the argument,

$$\begin{aligned} \bar{\mathbf{D}}_K: \mathbb{R}^{2k} &\rightarrow \mathbb{R}^{N \cdot n \cdot n}, \\ \bar{\mathbf{D}}_{K^T}: \mathbb{R}^{2k} &\rightarrow \mathbb{R}^{N \cdot n \cdot n}. \end{aligned} \quad (213)$$

Now, in the general form of the Lagrangian L in Eq. 11, the $m \times 2k$ matrix $\mathbf{A}$, with $m = Nn^2$, with N the number of grid points, and right-hand side m -vector $\mathbf{b}$, respectively, are given by

$$\begin{aligned} \mathbf{A} &= \bar{\mathbf{D}}_K + \bar{\mathbf{D}}_{K^T}, \\ \mathbf{b} &= 0. \end{aligned} \quad (214)$$

¹³Here, we can also see this directly, because for basis Killing fields $\mathbf{e}_i$, we have, as for all Killing fields, $\nabla \mathbf{e}_i + (\nabla \mathbf{e}_i)^T = 0$.

The $2k \times 2k$ Hessian of the Killing energy Lagrangian is thus

$$\mathbf{H}_L = (\bar{\mathbf{D}}_K + \bar{\mathbf{D}}_{K^T})^T (\bar{\mathbf{D}}_K + \bar{\mathbf{D}}_{K^T}). \quad (215)$$

However, the only non-zero $k \times k$ Hessian block is the top-left block

$$L_{\mathbf{q}\mathbf{q}} = (\mathbf{D}_K + \mathbf{D}_{K^T})^T (\mathbf{D}_K + \mathbf{D}_{K^T}). \quad (216)$$

Because we know that the Killing energy is independent of $\dot{\mathbf{q}}$, we immediately know that all other $k \times k$ Hessian blocks are zero, i.e.,

$$L_{\mathbf{q}\dot{\mathbf{q}}} = L_{\dot{\mathbf{q}}\mathbf{q}} = L_{\dot{\mathbf{q}}\dot{\mathbf{q}}} = 0. \quad (217)$$

M TENSOR AND NON-TENSOR QUANTITIES IN PHASE SPACE

The Lagrangian observer phase space $(\mathbf{q}, \dot{\mathbf{q}}, t) \cong \mathbb{R}^k \times \mathbb{R}^k \times \mathbb{R}$ has an interesting structure with several non-trivial implications. In the subsequent derivations, we follow the corresponding tensor calculus results given in the book by Lovelock and Rund [26]. Maybe somewhat surprisingly, the k -tuples of partial derivatives of the Lagrangian L with respect to the generalized coordinates q^i *do not form a tensor*:

$$\frac{\partial L(q^i, \dot{q}^i, t)}{\partial q^i} \text{ is not tensorial.} \quad (218)$$

The reason for this fact is the dependence of the Lagrangian L on the derivatives $\dot{\mathbf{q}}$,¹⁴ and the fact that the generalized velocities $\dot{\mathbf{q}}$ must be treated as tangent vectors to curves $t \mapsto \mathbf{q}(t)$, and must therefore transform accordingly to stay tangent to transformed curves $t \mapsto \tilde{\mathbf{q}}(t)$. Their transformation is determined by the transformation of the configuration space $\mathbf{q}$, and cannot be chosen independently, as derived below.

In contrast, the k -tuples of partial derivatives of the Lagrangian L with respect to the generalized velocities $\dot{q}^i$ *do form a covariant tensor*:

$$\frac{\partial L(q^i, \dot{q}^i, t)}{\partial \dot{q}^i} \text{ is a covariant vector.} \quad (219)$$

We can derive these properties of being tensorial or non-tensorial by considering an arbitrary change of coordinates. An arbitrary coordinate transformation of the phase space $(\mathbf{q}, \dot{\mathbf{q}}, t)$, independent of time t , i.e., identical for each time t , is only allowed to freely choose a diffeomorphism for the k generalized coordinates q^i , not for the k generalized velocities $\dot{q}^i$, because the latter are fully determined by the former. That is, for a change of generalized coordinates q^i given by¹⁵

$$\begin{aligned} \tilde{q}^i &= \tilde{q}^i(q^j) = \tilde{q}^i(q^1, \dots, q^k), \\ q^j &= q^j(\tilde{q}^i) = q^j(\tilde{q}^1, \dots, \tilde{q}^k), \end{aligned} \quad (220)$$

the corresponding generalized velocities $\dot{q}^i$ *must transform as*¹⁶

$$\dot{\tilde{q}}^i = \frac{\partial \tilde{q}^i}{\partial q^j} \dot{q}^j, \quad \dot{q}^j = \frac{\partial q^j}{\partial \tilde{q}^i} \dot{\tilde{q}}^i. \quad (221)$$

Now, however, since the Lagrangian L must be a scalar (with the tensorial meaning that L is a 0-order tensor field, i.e., the scalar field L is invariant to changes of coordinates), we must have

$$\begin{aligned} \tilde{L}(\tilde{q}^i, \dot{\tilde{q}}^i, t) &= L(q^j(\tilde{q}^i), \dot{q}^j(\tilde{q}^i, \dot{\tilde{q}}^i), t), \\ &= L(q^j(\tilde{q}^i), \frac{\partial q^j}{\partial \tilde{q}^i} \dot{\tilde{q}}^i, t). \end{aligned} \quad (222)$$

¹⁴We note that in degenerate cases where L does in fact not depend on $\dot{\mathbf{q}}$, i.e., $\partial L / \partial \dot{\mathbf{q}} = 0$ everywhere, this statement does not hold. For such degenerate Lagrangians L , the partial derivatives $\partial L / \partial q^i$ *do form a covariant vector*, as usual. But in these cases, there is no need for the phase space $(\mathbf{q}, \dot{\mathbf{q}}, t)$, and solutions to the Euler-Lagrange equation have to be computed in $(\mathbf{q}, t)$.

¹⁵The second row is the inverse of the transformation given by the first row.

¹⁶If the tangent vectors would not transform as given here, i.e., if they would not transform as contravariant vectors, the tangents of curves in phase space would not be independent of coordinates and would transform incorrectly.

From this invariance of the scalar field L , we can now directly derive the transformation laws of the partial derivatives of the Lagrangian L . Being non-tensorial means that the $\partial L/\partial q^i$ do not transform as a covariant (or contravariant) vector. This behavior is caused by the fact that in Eq. 222, the term $\dot{q}^j(\dots)$, which is the second argument to L on the right-hand side, depends not only on $\tilde{q}^i$ (which would be sufficient if the $\tilde{q}^i$ were independent variables with their own independent coordinate transformation, but here they are not independent variables), but also on $\tilde{q}^i$, because $\dot{q}^j = \tilde{q}^i \partial q^j / \partial \tilde{q}^i$: The coordinate Jacobian matrices $\partial q^j / \partial \tilde{q}^i$ depend on $\tilde{q}^i$.¹⁷ Now, by differentiating Eq. 222 with respect to $\tilde{q}^i$, and using that from Eq. 221 (right equation) we can obtain

$$\frac{\partial \dot{q}^j}{\partial \tilde{q}^i} = \frac{\partial^2 q^j}{\partial \tilde{q}^i \partial \tilde{q}^k} \dot{q}^k, \quad (223)$$

we can derive that the $\partial L/\partial q^i$ transform as

$$\begin{aligned} \frac{\partial \tilde{L}}{\partial \tilde{q}^i} &= \frac{\partial q^j}{\partial \tilde{q}^i} \frac{\partial L}{\partial q^j} + \frac{\partial \dot{q}^j}{\partial \tilde{q}^i} \frac{\partial L}{\partial \dot{q}^j}, \\ &= \frac{\partial q^j}{\partial \tilde{q}^i} \frac{\partial L}{\partial q^j} + \frac{\partial^2 q^j}{\partial \tilde{q}^i \partial \tilde{q}^k} \dot{q}^k \frac{\partial L}{\partial \dot{q}^j}. \end{aligned} \quad (224)$$

Thus, because of the second term they do not form a tensor. In contrast, by differentiating Eq. 222 with respect to $\tilde{q}^i$, and using that from Eq. 220 (second row) and Eq. 221 (right equation) we can, respectively, obtain

$$\frac{\partial q^j}{\partial \tilde{q}^i} = 0, \quad \frac{\partial \dot{q}^j}{\partial \tilde{q}^i} = \frac{\partial q^j}{\partial \tilde{q}^i}, \quad (225)$$

we can derive that the partial derivatives $\partial L/\partial \dot{q}^i$ do transform as a covariant vector and are thus a tensor, because they transform as

$$\begin{aligned} \frac{\partial \tilde{L}}{\partial \dot{q}^i} &= \frac{\partial q^j}{\partial \tilde{q}^i} \frac{\partial L}{\partial q^j} + \frac{\partial \dot{q}^j}{\partial \tilde{q}^i} \frac{\partial L}{\partial \dot{q}^j}, \\ &= \frac{\partial q^j}{\partial \tilde{q}^i} \frac{\partial L}{\partial q^j}. \end{aligned} \quad (226)$$

M.1 Tensor Property of the Euler-Lagrange Operator

It is a crucial property of the calculus of variations that the Euler-Lagrange operator ψ , written in coordinate components as

$$\psi_i = \frac{\partial L}{\partial q^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i}, \quad (227)$$

is a covariant vector [26]. That is, it indeed transforms as

$$\tilde{\psi}_i = \frac{\partial q^j}{\partial \tilde{q}^i} \psi_j. \quad (228)$$

We can also write the Euler-Lagrange operator ψ_i explicitly as

$$\psi_i(t) = \frac{\partial L}{\partial q^i} - \frac{\partial^2 L}{\partial \dot{q}^i \partial q^j} \dot{q}^j(t) - \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \ddot{q}^j(t) - \frac{\partial^2 L}{\partial t \partial \dot{q}^i}. \quad (229)$$

To derive the transformation law of the Euler-Lagrange operator, we can first compute the time derivative (i.e., derivative in direction $\dot{\mathbf{q}}$)

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial \tilde{L}}{\partial \dot{q}^i} \right) &= \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^j} \right) \frac{\partial q^j}{\partial \tilde{q}^i} + \frac{\partial L}{\partial \dot{q}^j} \frac{d}{dt} \left(\frac{\partial q^j}{\partial \tilde{q}^i} \right), \\ &= \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^j} \right) \frac{\partial q^j}{\partial \tilde{q}^i} + \frac{\partial L}{\partial \dot{q}^j} \frac{\partial^2 q^j}{\partial \tilde{q}^i \partial \tilde{q}^k} \dot{q}^k. \end{aligned} \quad (230)$$

¹⁷And on q^i , but these coordinates are determined directly by the $\tilde{q}^i$ via $q^j = q^j(\tilde{q}^i)$, and therefore knowing the coordinates $\tilde{q}^i$, $\tilde{q}^i$ is sufficient.

If we now subtract each side of the above from each side, respectively, of the transformation given by Eq. 224, we obtain the transformation

$$\frac{\partial \tilde{L}}{\partial \tilde{q}^i} - \frac{d}{dt} \left(\frac{\partial \tilde{L}}{\partial \dot{q}^i} \right) = \frac{\partial q^j}{\partial \tilde{q}^i} \left(\frac{\partial L}{\partial q^j} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^j} \right) \right). \quad (231)$$

This now proves that the Euler-Lagrange operator ψ_j (right-hand side) transforms as a covariant vector to $\tilde{\psi}_i$ (left-hand side), as in Eq. 228.

M.2 Tensor Property for Curves Defined Point-Wise

Considering that the partial derivatives $L_{\mathbf{q}} = \partial L/\partial \mathbf{q}$ do not form a tensor, an important question is:

If we have an observer given by a curve $t \mapsto \mathbf{q}(t)$ defined point-wise using Eq. 34, could it be that in some other coordinate system $\partial L/\partial \mathbf{q}$ would actually be non-zero somewhere on the curve?

Fortunately, the answer is *no*: From the actual (non-tensorial) transformation law of the partial derivatives $L_{\mathbf{q}}$ (Eq. 224), we see that at all points where the tensor $\mathbf{p} = \partial L/\partial \dot{\mathbf{q}}$ vanishes, and only there, the transformation law of $L_{\mathbf{q}}$ is identical to the transformation law of a covariant tensor. This has the crucial consequence that, if $\mathbf{p} = \partial L/\partial \dot{\mathbf{q}} = 0$, if $L_{\mathbf{q}} = 0$ in some (any) coordinate system, we are guaranteed to have $L_{\mathbf{q}} = 0$ in all other coordinate systems.