CrossMark

Variational integrators for nonvariational partial differential equations

Michael Kraus [a,b,*], Omar Maj [a]

[a] Max-Planck-Institut für Plasmaphysik, Boltzmannstraße 2, 85748 Garching, Germany
[b] Technische Universität München, Zentrum Mathematik, Boltzmannstraße 3, 85748 Garching, Germany

HIGHLIGHTS

- Combination of Ibragimov's theory of formal Lagrangians with variational integrators.
- Significant extension of the applicability of variational integrators.
- Numerical schemes for the 2D vorticity equation preserving circulation, enstrophy and energy.
- Covariant generalisation of Arakawa's discretisation of the Jacobian.

ARTICLE INFO

ABSTRACT

Variational integrators for Lagrangian dynamical systems provide a systematic way to derive geometric numerical methods. These methods preserve a discrete multisymplectic form as well as momenta associated to symmetries of the Lagrangian via Noether's theorem. An inevitable prerequisite for the derivation of variational integrators is the existence of a variational formulation for the considered problem. Even though for a large class of systems this requirement is fulfilled, there are many interesting examples which do not belong to this class, e.g., equations of advection–diffusion type frequently encountered in fluid dynamics or plasma physics.

On the other hand, it is always possible to embed an arbitrary dynamical system into a larger Lagrangian system using the method of formal (or adjoint) Lagrangians. We investigate the application of the variational integrator method to formal Lagrangians, and thereby extend the application domain of variational integrators to include potentially all dynamical systems.

The theory is supported by physically relevant examples, such as the advection equation and the vorticity equation, and numerically verified. Remarkably, the integrator for the vorticity equation combines Arakawa's discretisation of the Poisson brackets with a symplectic time stepping scheme in a fully covariant way such that the discrete energy is exactly preserved. In the presentation of the results, we try to make the geometric framework of variational integrators accessible to non specialists.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

In recent years, the field of structure-preserving or geometric discretisation [1–3] has become a flourishing discipline of numerical analysis and scientific computing. One particular family of geometric discretisation methods is that of variational integrators [4–10], which are based on the discretisation of Hamilton's principle of stationary action [11–15]. Variational integrators preserve a discrete multisymplectic form and have good longtime energy behaviour. As we will see, they can be designed to preserve energy even exactly, which in practice means up to machine precision. Furthermore, they preserve momenta associated to symmetries of the discrete equations of motion via a discrete version of Noether's theorem [16,17].

While in most standard discretisation techniques for dynamical systems the equations of motion are directly discretised, the basic idea of variational integrators is to construct a discrete counterpart to the considered system. This means that the fundamental building blocks

of classical mechanics and field theory, namely the action functional, the Lagrangian, the variational principle, and the Noether theorem, all have discrete equivalents. The application of the discrete variational principle to the discrete action then leads to discrete Euler–Lagrange equations. The evolution map that corresponds to the discrete Euler–Lagrange equations is what is called a variational integrator. The discrete Noether theorem can be used to relate symmetries of the discretised system to discrete momenta that are exactly preserved by this integrator. Whereas most standard techniques put emphasis on the minimisation of local errors, for variational integrators the focus is rather on the preservation of global or geometric properties of the system.

An obvious limitation of the variational integrator method is its applicability to Lagrangian systems only. This excludes a large class of interesting systems, for example the problems of advection–diffusion type often found in fluid dynamics and plasma physics. We propose here that the method of formal (or adjoint) Lagrangians [18] can be used as an expedient to avoid this limitation. More specifically, formal Lagrangians allow us to embed any given system into a larger system which, in turn, admits a Lagrangian formulation. To obtain a formal Lagrangian L, the equation at hand, say $\mathcal{F}[u] = 0$, is multiplied by an adjoint variable v, giving $L = v \cdot \mathcal{F}[u]$. Variation of the resulting action functional, $\mathcal{A} = \int L \, d^{n+1}x$, with respect to v gives the original equation $\mathcal{F}[u] = 0$. Variation of the action functional with respect to the physical variable u gives an additional equation that determines the evolution of the adjoint variable v.

At first sight one might be tempted to regard the formal Lagrangian formalism as merely a method for obtaining a weak formulation of the problem at hand. Then, if our goal is to obtain an integrator, the details of the dynamics of the adjoint variable v would seem irrelevant. However, it turns out that the dynamics of v play a key role in relating symmetries of the formal Lagrangian to conservation laws satisfied by u. Ibragimov [19–21] developed a theory for the analysis of conservation laws of arbitrary differential equations by extending the Noether theorem to formal Lagrangians. This leads to conservation laws for the extended system (u, v), which can be restricted to the original system provided that it is possible to express the solution of the adjoint variable v in terms of u.

In this work, we propose the combination of the discrete variational principle with Ibragimov's theory in order to derive variational integrators for systems without a natural Lagrangian formulation and to determine the associated discrete conservation laws. Thereby we extend significantly the applicability of the variational integrator method. The goal of this approach is to design numerical schemes which respect certain conservation laws of a given system in a rather systematic way.

We proceed as follows. In Section 2, we present the theory of variational integrators in simple terminology. To set the stage and fix notation we review the continuous action principle for field theories and the corresponding Noether theorem before passing over to the discrete theory, which is extended to account for discrete divergence symmetries. The style of presentation is chosen to make the theory accessible to a wide audience without extensive background in modern differential geometry. This implies some loss of generality, but hopefully not too much of the geometric beauty of the original work is lost. In Section 3, we recall the inverse problem of the calculus of variations, review the theory of formal Lagrangians and explain the derivation of conserved quantities in this setting. We also provide a geometric formulation of the theory, which to our knowledge has not been presented, yet. Finally, in Section 4, we apply the method to some prototypical examples, including the advection equation and the vorticity equation, and verify the theoretical properties in numerical experiments. More elaborate numerical examples for the Vlasov–Poisson system as well as ideal and reduced magnetohydrodynamics will be presented elsewhere [22–24]. The examples we provide here can be seen as building blocks for these more complicated systems.

2. Variational integrators

2.1. Geometry and notation

In this work, we are concerned with the discretisation of partial differential equations (PDEs) of evolution type. A field is a map $u : X \to F$ from a bounded domain $X \subset \mathbb{R}^{n+1}$ taking values in an open set $F \subseteq \mathbb{R}^m$. Most often, X corresponds to some region of spacetime with coordinates

$$\mathrm{x} = (\mathrm{x}^\mu) = (\mathrm{x}^0, \mathrm{x}^i) = (t, x, y, z) \quad \text{with } 0 \le \mu \le n, \ 1 \le i \le n,$$

and $n = \dim X - 1$ being the number of space-like dimensions. Most of the theoretic results rely neither on this space-plus-time splitting nor on X being a subset of an Euclidean space. Thus X can be replaced by a differentiable manifold, in which case the results are valid locally in a coordinate chart. Points on F are denoted by $\mathrm{y} = (\mathrm{y}^a)$ with $1 \le a \le m$ and $m = \dim F$ being the number of field components. We make use of the Einstein summation convention on repeated indices, both for coordinates and fields.

The configuration of a field u is geometrically represented by its graph

$$\mathrm{graph}(u) = \big\{ (\mathrm{x}, \mathrm{y}) \,\big|\, \mathrm{y} = u(\mathrm{x}) \big\},$$

which is a subset of the Cartesian product

$$Y = X \times F = \big\{ (\mathrm{x}, \mathrm{y}) \,\big|\, \mathrm{x} \in X, \ \mathrm{y} \in F \big\}.$$

With the projection onto the first factor,

$$\begin{aligned} \pi \ : \quad Y \quad &\to \quad X, \\ (\mathrm{x}, \mathrm{y}) \quad &\mapsto \quad \mathrm{x}, \end{aligned}$$

we can define a geometrical structure (Y, X, π) called (trivial) fibre bundle. Here, X is called the base space, Y the configuration space, and F the fibre. Local coordinates on Y are given by

$$(\mathrm{x}^\mu, \mathrm{y}^a) \quad \text{with } 0 \le \mu \le n, \ 1 \le a \le m.$$

A field u can be identified with a section of the bundle, i.e., a map $\varphi : X \to Y$, satisfying the condition

$$\pi \circ \varphi = \mathrm{id}_X,$$

where id$_X$ is the identity map on X. This condition ensures that the image of a section φ corresponds to the graph of a field u,

$$\varphi(X) = \big\{(x, y) = \varphi(x) \mid x \in X\big\} = \big\{(x, u(x)) \mid x \in X\big\}.$$

In local coordinates, a section $\varphi : X \to Y$ can be written as $\varphi(x) = (x^\mu, \varphi^a(x))$, that is $\varphi(x)$ is given by functions $y^a = \varphi^a(x) = u^a(x)$.

Partial derivatives of a field component u^a with respect to x^μ or x^i are denoted u^a_μ or u^a_i, respectively. The collection of all partial derivatives of a given order k is denoted $u_{(k)}$. The appropriate geometric setting for partial derivatives of a field u is the theory of jet bundles [25–29]. Jets combine into a single object the values of a field at a point and the values of its derivatives. More specifically, the jet prolongation of a section φ is a map

$$j^1\varphi : x \mapsto \big(x^\mu, \varphi^a(x), \varphi^a_\mu(x)\big),$$

taking values in $Y \times \mathbb{R}^{m(n+1)}$. This space is identified with the first jet bundle of Y, denoted J^1Y, so that the first jet prolongation of a field is a section of J^1Y. Local coordinates on J^1Y are given by

$$(x^\mu, y^a, z^a_\mu) \quad \text{with } 0 \le \mu \le n, \ 1 \le a \le m,$$

where z^a_μ represents the possible values of partial derivatives. The Lagrangian of a first order field theory for example will be a function defined on the first jet bundle J^1Y. The jet bundle J^1Y has a natural projection on both Y and X, denoted by π_Y and π_X, respectively. The latter defines a fibre bundle (J^1Y, X, π_X). It is worth noting that not every section ψ of the bundle $J^1Y \to X$ is the jet prolongation of a section $\varphi : X \to Y$. When that happens we say that ψ is holonomic.

Higher order jet bundles J^kY of Y can be defined analogously by considering the kth jet prolongation of a section φ,

$$j^k\varphi : x \mapsto \big(x^\mu, \varphi^a(x), \varphi^a_\mu(x), \varphi^a_{\mu\nu}(x), \ldots\big),$$

which includes derivatives of φ up to kth order. Higher-order jet bundles J^kY can also be defined as submanifolds of iterated first-order jet bundles, e.g., J^2Y is a submanifold of $J^1(J^1Y)$. If coordinates on $J^1(J^1Y)$ are denoted by

$$(x^\mu, y^a, z^a_\mu, \bar{z}^a_\mu, w^a_{\mu\nu}) \quad \text{with } 0 \le \mu, \nu \le n, \ 1 \le a \le m,$$

then J^2Y is obtained by requiring that $z^a_\mu = \bar{z}^a_\mu$ for all a and all μ, and accordingly for larger k (for more details see e.g. [26, Chapter 32] or [29, Section 4.1.4]).

Jet bundles have become the standard framework for PDE analysis and variational calculus. They provide a natural setting for the formulation of field theories and the analysis of conservation laws which are central to this work. This framework generalises in a standard way to non-trivial cases in which X is a manifold and Y cannot be written globally as a Cartesian product. For a thorough introduction into the geometric framework we refer to Gotay et al. [30].

2.2. Continuous action principle

For definiteness, let us consider only first order Lagrangian field theories, i.e., the Lagrangian L shall depend only on the coordinates, the fields and their first derivatives. Therefore, the Lagrangian is a function defined on the first jet bundle,

$$L : J^1Y \to \mathbb{R}. \tag{1}$$

The corresponding action functional is

$$\mathcal{A}[\varphi] = \int_X L\big(j^1\varphi\big)\, d^{n+1}x. \tag{2}$$

Hamilton's principle of stationary action [11,12,15] states that among all possible field configurations $\varphi : X \to Y$, the one chosen by nature makes the action functional (2) stationary. As usual in geometric mechanics [13,14,30,6], stationary points of the action are meant in the formal sense. The variations of a field φ are defined in terms of a geometric transformation of the underlying bundle Y, that is a sufficiently regular one-parameter group of transformations $\sigma^\epsilon(x, y)$ of Y into itself, defined for ϵ in a neighbourhood of zero. In order to have near identity transformations, it is required that σ^ϵ reduces to the identity map at $\epsilon = 0$, i.e., $\sigma^\epsilon|_{\epsilon=0} = \text{id}$. Furthermore, variations should vanish at boundary points, so that $\sigma^\epsilon = \text{id}$ for $x \in \partial X$. The map $\sigma^\epsilon(x, y)$ is the flow of the vector field V over Y given by

$$V(x, y) = \frac{d}{d\epsilon}\sigma^\epsilon(x, y)\bigg|_{\epsilon=0}, \tag{3}$$

and written in components as

$$V(x, y) = \eta^a(x, y)\, \frac{\partial}{\partial y^a}, \tag{4}$$

where we have $\eta^a(x, y) = 0$ for $x \in \partial X$. Here, vector fields are identified with first-order differential operators. The variation of a field φ in the direction of V is then defined by

$$\varphi^\epsilon = (\sigma^\epsilon \circ \varphi)(x) = \sigma^\epsilon(\varphi(x)), \tag{5}$$

and we have $\varphi^\epsilon|_{\epsilon=0} = \varphi$ and $\varphi^\epsilon = \varphi$ on boundary points $x \in \partial X$. Loosely speaking, φ^ϵ is obtained by dragging φ along the flow of the vector field V. For sake of simplicity, we consider transformations in σ^ϵ that leave the point x unchanged, thus, preserving the fibres of the

bundle Y. Those are referred to as vertical transformations.[1] We say that φ is a stationary point of $\mathcal{A}[\varphi]$ if for every σ^ϵ we find that $\epsilon = 0$ is a stationary point of $\mathcal{A}[\varphi^\epsilon]$ viewed as a function of ϵ in an open interval about $\epsilon = 0$. Then Hamilton's principle amounts to

$$\frac{d}{d\epsilon} \mathcal{A}[\sigma^\epsilon \circ \varphi]\Big|_{\epsilon=0} = 0 \quad \text{for every } \sigma^\epsilon. \tag{6}$$

Explicitly,

$$\frac{d}{d\epsilon} \mathcal{A}[\sigma^\epsilon \circ \varphi]\Big|_{\epsilon=0} = \int_X \frac{d}{d\epsilon}\Big[L\big(j^1(\sigma^\epsilon \circ \varphi)\big)\Big]\Big|_{\epsilon=0} d^{n+1}x \tag{7}$$

$$= \int_X \left[\frac{\partial L}{\partial y^a}\big(j^1\varphi\big) \cdot \frac{d(\sigma^\epsilon)^a}{d\epsilon} + \frac{\partial L}{\partial z_\mu^a}\big(j^1\varphi\big) \cdot \frac{d}{d\epsilon}\frac{\partial(\sigma^\epsilon)^a}{\partial x^\mu} \right]\Big|_{\epsilon=0} d^{n+1}x. \tag{8}$$

The second term in (8) can be integrated by parts,

$$\frac{d}{d\epsilon} \mathcal{A}[\sigma^\epsilon \circ \varphi]\Big|_{\epsilon=0} = \int_X \left[\frac{\partial L}{\partial y^a}\big(j^1\varphi\big) - \frac{\partial}{\partial x^\mu}\left(\frac{\partial L}{\partial z_\mu^a}\big(j^1\varphi\big) \right) \right] \cdot \eta^a(\varphi)\, d^{n+1}x, \tag{9}$$

where we used (3) and (4). As the variation of the action has to vanish for arbitrary transformations σ^ϵ and therefore for arbitrary vector fields V, the term in square brackets has to vanish identically, that is

$$\big(j^2\varphi\big)^*\big(D_{EL}(L)\big)^a = \frac{\partial L}{\partial y^a}\big(j^1\varphi\big) - \frac{\partial}{\partial x^\mu}\left(\frac{\partial L}{\partial z_\mu^a}\big(j^1\varphi\big) \right) = 0. \tag{10}$$

These are the Euler–Lagrange field equations, i.e., the equations of motion for a first order Lagrangian field theory, which for a non-degenerate Lagrangian L live in the second-order jet bundle $J^2 Y$. By D_{EL} we denote the Euler–Lagrange operator acting on the Lagrangian L. The pull-back notation $(j^2\varphi)^*$ states that the result is evaluated on the second jet prolongation $j^2\varphi$. The theory is fully covariant, that is for time-dependent problems, time is regarded as a component of x.

2.3. Continuous Noether theorem

The Noether theorem [16,17] states that each Lie point symmetry of a Lagrangian corresponds to a conservation law of the associated Euler–Lagrange equations.

We restrict our attention to conservation laws that are generated by vertical transformations of the configuration bundle Y, i.e., transformations which leave the base space X invariant. In the framework of formal Lagrangians addressed below, this is often sufficient to uncover interesting conservation laws (including conservation of momentum and energy). Our derivation of the Noether theorem is essentially based on reference [6].

As in Section 2.2, a transformation is generated by a map $\sigma^\epsilon(x, y)$, that is

$$\varphi^\epsilon = \sigma^\epsilon \circ \varphi \quad \text{with } \sigma^\epsilon|_{\epsilon=0} = \text{id and } V = \frac{d\sigma^\epsilon}{d\epsilon}\Big|_{\epsilon=0}, \tag{11}$$

but it is not required that σ^ϵ reduces to the identity at boundary points. In our analysis we will usually just prescribe the generating vector field V instead of the actual transformation σ^ϵ. In components, V can be written as

$$V(x, y) = \eta^a(x, y)\frac{\partial}{\partial y^a} \quad \text{with } \eta^a(x, y) = \frac{d}{d\epsilon}(\sigma^\epsilon)^a(x, y)\Big|_{\epsilon=0}, \tag{12}$$

where $(\sigma^\epsilon)^a$ is the ath component of the transformation map. Since Lagrangians are functions defined on $J^1 Y$, we need to compute the first prolongation of the generating vector field. The prolongation of V is defined via the jet prolongation of the transformation map σ^ϵ,

$$j^1 V = \frac{d}{d\epsilon}\Big[j^1\sigma^\epsilon(x, y)\Big]\Big|_{\epsilon=0}, \tag{13}$$

and is given by

$$j^1 V = \eta^a\frac{\partial}{\partial y^a} + \left(\frac{\partial\eta^a}{\partial x^\mu} + z_\mu^b\frac{\partial\eta^a}{\partial y^b} \right)\frac{\partial}{\partial z_\mu^a} = \eta^a\frac{\partial}{\partial y^a} + \eta_\mu^a\frac{\partial}{\partial z_\mu^a}. \tag{14}$$

A vertical transformation σ^ϵ is a symmetry transformation for the Lagrangian (1) if the invariance condition,

$$L\big(j^1(\sigma^\epsilon \circ \varphi)\big) = L\big(j^1\varphi\big), \tag{15}$$

---

[1] As noted in [6,31], one really should consider general transformations that might affect the points of the base space as well as the fields. Otherwise, the Cartan form and thus the multisymplectic form will not be correctly obtained from the variational principle. However, in this work we are only concerned with the Euler–Lagrange equations and the Noether theorem. Both will be obtained correctly even if only transformations of the fields (vertical transformations in the language of jet bundles) are considered. Transformations of the base space (horizontal transformations) would substantially complicate the derivations and are neglected for sake of simplicity.

is satisfied.[2] Taking the ϵ derivative of (15), we obtain an infinitesimal invariance condition,

$$\frac{d}{d\epsilon} L\big(j^1(\sigma^\epsilon \circ \varphi)\big)\bigg|_{\epsilon=0} = 0, \tag{16}$$

which is equivalent to (15). Explicitly computing the ϵ derivative, we obtain

$$j^1 V(L)(j^1\varphi) = \frac{\partial L}{\partial y^a}(j^1\varphi) \cdot \eta^a(\varphi) + \frac{\partial L}{\partial z_\mu^a}(j^1\varphi) \cdot \left[\frac{\partial \eta^a}{\partial x^\mu} + \frac{\partial \eta^a}{\partial y^b}\frac{\partial \varphi^b}{\partial x^\mu}\right] = 0. \tag{17}$$

If φ solves the Euler–Lagrange field equations (10), we can replace the first term on the right-hand side of (17) to obtain

$$\left[\frac{\partial}{\partial x^\mu}\frac{\partial L}{\partial z_\mu^a}(j^1\varphi)\right] \cdot \eta^a(\varphi) + \frac{\partial L}{\partial z_\mu^a}(j^1\varphi) \cdot \left[\frac{\partial \eta^a}{\partial x^\mu} + \frac{\partial \eta^a}{\partial y^b}\frac{\partial \varphi^b}{\partial x^\mu}\right] = 0. \tag{18}$$

This, at last, amounts to a total divergence,

$$\mathrm{div}\big[J(j^1\varphi)\big] = 0, \tag{19}$$

with the Noether current J given by

$$J^\mu\big(j^1\varphi\big) = \frac{\partial L}{\partial z_\mu^a}(j^1\varphi) \cdot \eta^a(\varphi). \tag{20}$$

The fact that the Noether current is divergence-free, expresses the conservation law satisfied by solutions φ of the Euler–Lagrange field equation (10). The flux of J through the boundary of any domain $\Omega \subseteq X$ is zero.

Global form of conservation laws

Let us consider an alternative point of view that is better suited for actual computations on the discrete level. If the Lagrangian is invariant under the vertical transformation (11), the action is also invariant under this transformation,

$$\frac{d}{d\epsilon}\mathcal{A}[\sigma^\epsilon \circ \varphi]\bigg|_{\epsilon=0} = \int_X j^1 V(L)(j^1\varphi)\, d^{n+1}x = 0. \tag{21}$$

Repeating the steps that lead from (17) to (19) this becomes

$$\frac{d}{d\epsilon}\mathcal{A}[\sigma^\epsilon \circ \varphi]\bigg|_{\epsilon=0} = \int_X \frac{\partial}{\partial x^\mu}\left[\frac{\partial L}{\partial z_\mu^a}(j^1\varphi) \cdot \eta^a(\varphi)\right] d^{n+1}x = 0. \tag{22}$$

With appropriate boundary conditions, one has

$$\int \frac{\partial}{\partial x^i}\left[\frac{\partial L}{\partial z_i^a}(j^1\varphi) \cdot \eta^a(\varphi)\right] d^n x = 0, \tag{23}$$

and thus, integrating (22) for $x^0 = t \in [t_0, t_1]$,

$$\frac{d}{d\epsilon}\mathcal{A}[\sigma^\epsilon \circ \varphi]\bigg|_{\epsilon=0} = \left[\int \frac{\partial L}{\partial z_t^a}(j^1\varphi) \cdot \eta^a(\varphi)\, d^n x\right]_{t_0}^{t_1} = 0. \tag{24}$$

This is equivalent to integrating the divergence of the Noether current (19) over the spatial dimensions and using (23). Since t_0 and t_1 are arbitrary, this implies the conservation of

$$\mathcal{J} = \int \frac{\partial L}{\partial z_t^a}(j^1\varphi) \cdot \eta^a(\varphi)\, d^n x = \int J^t(j^1\varphi)\, d^n x, \tag{25}$$

which is called Noether charge.

Divergence symmetries

Sometimes it is necessary to consider a slightly more general version of Noether's theorem [33,20]. The invariance condition (17) can be weakened to

$$j^1 V(L)(j^1\varphi) = \mathrm{div}\, B(j^1\varphi), \tag{26}$$

with B being a vector field over Y of the form $B = B^\mu(x, y)\, \partial/\partial x^\mu$, and $\mathrm{div}\, B$ is a scalar field defined over $J^1 Y$ given by

$$\mathrm{div}\, B(x, y, z) = \frac{\partial B^\mu(x, y)}{\partial x^\mu} + \frac{\partial B^\mu(x, y)}{\partial y^a} z_\mu^a. \tag{27}$$

This form of the invariance condition is tightly related to the gauge freedom of the Lagrangian. We can add any divergence to the Lagrangian without changing the equations of motion as

---

[2] Here, invariance of the Lagrangian suffices, but for general transformations, i.e., transformations that transform the base space in addition to the configuration space, invariance of the Lagrangian does not guarantee that φ^ϵ is a solution whenever φ is a solution. Invariance of solutions requires invariance of the action, and invariance of the action requires equivariance of the Lagrangian, which also takes the deformation of the integration domain X into account. For a detailed discussion see, e.g., Lew et al. [32], Section 6.5.
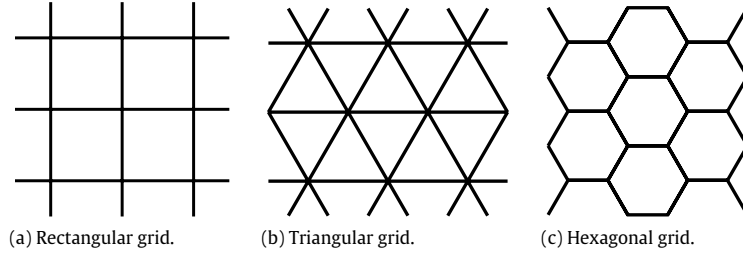
(a) Rectangular grid.   (b) Triangular grid.   (c) Hexagonal grid.

Fig. 1. Different regular grids.

$$\mathcal{A}[\varphi] = \int_X L'(\mathrm{j}^1\varphi)\, d^{n+1}\mathrm{x} = \int_X \big(L + \operatorname{div} H\big)(\mathrm{j}^1\varphi)\, d^{n+1}\mathrm{x} = \int_X L(\mathrm{j}^1\varphi)\, d^{n+1}\mathrm{x}, \tag{28}$$

assuming appropriate boundary conditions. Here, H is a vector field over Y of the form $H = H^\mu(\mathrm{x}, \mathrm{y})\, \partial/\partial x^\mu$, $\operatorname{div} H$ is defined as in (27), and

$$L'(\mathrm{x}, \mathrm{y}, \mathrm{z}) = \big(L + \operatorname{div} H\big)(\mathrm{x}, \mathrm{y}, \mathrm{z}). \tag{29}$$

One can check that the Euler–Lagrange equations for L and L' are the same. Requiring the invariance condition (17) to be satisfied for L',

$$\mathrm{j}^1 V(L')(\mathrm{j}^1\varphi) = \mathrm{j}^1 V(L)(\mathrm{j}^1\varphi) + \operatorname{div} \tilde{H}(\mathrm{j}^1\varphi) = 0, \quad \text{with } \tilde{H} = V(H^\mu)\frac{\partial}{\partial x^\mu}, \tag{30}$$

we find the following relation between the vector fields B and H,

$$B = -\tilde{H}. \tag{31}$$

In summary, if V is a divergence symmetry for L in the sense of (26) and there exists a vector field H satisfying (31), then V is a Lie-point symmetry (17) of the equivalent Lagrangian L' defined in (29). We can therefore apply the Noether theorem to L', obtaining the Noether current

$$J'(\mathrm{j}^1\varphi) = J(\mathrm{j}^1\varphi) - B(\varphi), \tag{32}$$

with J' the conserved Noether current associated to the divergence symmetry (26) and J as defined in (20), which is not conserved in this case.

2.4. Discrete jet bundles

In order to derive a discrete version of Hamilton's action principle, we proceed in three steps. In this section, we define discrete analogues of the base space X, the configuration space Y, and the jet bundles $\mathrm{J}^1 Y$ and $\mathrm{J}^2 Y$. In the next section, we discretise the Lagrangian and the action functional, and finally, we work out the discrete action principle.

The continuous base space X is replaced by its discrete analogue X_d, a bounded subset $X_d \subset \mathbb{Z}^{n+1}$ which corresponds to a grid of points in X. In what follows we assume X to be two-dimensional and choose an equidistant rectangular discretisation like it is depicted in Fig. 1(a). The theory is easily applicable to other types of grids as well, e.g., triangular as in Fig. 1(b), hexagonal as in Fig. 1(c), or even staggered or irregular ones.

The coordinates on the discrete base space $X_d \subset \mathbb{Z} \times \mathbb{Z}$ are denoted (i, j) and assumed to take values $i \in \{1, \ldots, N_0\}$ and $j \in \{1, \ldots, N_1\}$, respectively, where N_0 and N_1 are the number of points for each dimension. This corresponds to a grid of points $\mathrm{x}_{i,j} = (ih, jh)$ in X with, for simplicity, the same step size h in both directions. Then

$$X_d \cong \big\{ (ih, jh) \in X \mid i = 1, \ldots, N_0,\ j = 1, \ldots, N_1 \big\}. \tag{33}$$

The discrete configuration space is defined as the cartesian product

$$Y_d = X_d \times F, \tag{34}$$

where F is the same as in the continuous case and we have an analogous projection

$$\begin{aligned}
\pi_d \quad : \quad Y_d \quad &\rightarrow \quad X_d, \\
(\mathrm{y}_{i,j}) \quad &\mapsto \quad (i, j).
\end{aligned} \tag{35}$$

Coordinates of Y_d are denoted $\mathrm{y}_{i,j}^a$ with $1 \le a \le \dim F$. The coordinates of the base point (i, j) are already implied and therefore not specified separately. While coordinates on X_d and Y_d are defined point-wise, coordinates on $\mathrm{J}^1 Y_d$ will be defined grid-cell-wise. We therefore introduce the following abstract but convenient notation. A square $\square$ on X_d is an ordered quadruplet

$$\square = \big((i, j), (i, j+1), (i+1, j+1), (i+1, j) \big), \tag{36}$$

defining a primal grid cell. The set of such cells on X_d is denoted $X^\square$. Vertices $\square^l$ of a square $\square$ with $1 \le l \le 4$ are counted counter-clockwise from the bottom left (c.f. Fig. 2(a)), namely,

$$\square^1 = (i, j), \qquad \square^2 = (i, j+1), \qquad \square^3 = (i+1, j+1), \qquad \square^4 = (i+1, j). \tag{37}$$

(a) Grid points.
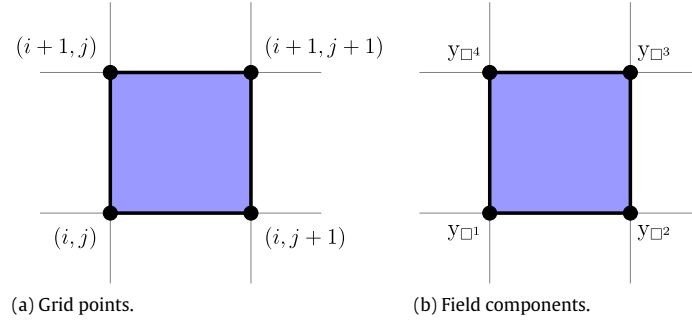
(b) Field components.

Fig. 2. Vertices of a primal grid cell in a two-dimensional rectangular grid and field components at those vertices.

On a quadrilateral grid, the first jet bundle of Y_d becomes [6]

$$J^1 Y_d = X^\square \times F^4, \tag{38}$$

with coordinates given by

$$\left\{ (\square, y^a_{\square l}) \,\middle|\, \square \in X^\square, \ y^a_{\square l} \in \mathbb{R}, \ 1 \le l \le 4, \ 1 \le a \le \dim F \right\}. \tag{39}$$

This implies that if $\square^1 = (i, j)$, then we have

$$y_{\square^1} = y_{i,j}, \qquad y_{\square^2} = y_{i,j+1}, \qquad y_{\square^3} = y_{i+1,j+1}, \qquad y_{\square^4} = y_{i+1,j}. \tag{40}$$

A section of Y_d, representing a discrete field, is a map $\varphi_d : X_d \to Y_d$ such that $\pi \circ \varphi_d = \mathrm{id}_{X_d}$. The components of a discrete field φ_d at the vertices (37) of a specific square are

$$\varphi_{\square^1} = \varphi_d(\square^1), \qquad \varphi_{\square^2} = \varphi_d(\square^2), \qquad \varphi_{\square^3} = \varphi_d(\square^3), \qquad \varphi_{\square^4} = \varphi_d(\square^4), \tag{41}$$

as depicted in Fig. 2(b). The discrete first jet prolongation of a discrete field φ_d is a map $j^1 \varphi_d : X^\square \to J^1 Y_d$, defined as

$$j^1 \varphi_d(\square) = \left(\square, \varphi_{\square^1}, \varphi_{\square^2}, \varphi_{\square^3}, \varphi_{\square^4} \right). \tag{42}$$

This contains all the information necessary to define discrete first-order derivatives.

We will see that the discrete Euler–Lagrange equations are living on $J^2 Y_d$, i.e., the second jet bundle of Y_d, similar to the continuous case (see [7] for another discussion of discrete second order jet bundles). Coordinates on $J^2 Y_d$ will be defined on a quadruple of cells $\boxplus$, namely those cells $\square$ in $X^\square$ which share a common vertex. The cells $\square_m$ of $\boxplus$ with $1 \le m \le 4$ are counted counter-clockwise from the bottom left, so that

$$\boxplus = \left\{ \square_1, \square_2, \square_3, \square_4 \in X^\square \,\middle|\, \square_1^3 = \square_2^4 = \square_3^1 = \square_4^2 \right\}. \tag{43}$$

The set of all such quadruples on X_d is denoted by $X^\boxplus$. Vertices $\boxplus^l$ of a quadruple $\boxplus$ with $1 \le l \le 9$ are counted row-wise from bottom left to top right, namely,

$$\boxplus^1 = (i-1, j-1), \qquad\qquad \boxplus^2 = (i-1, j), \qquad\qquad \boxplus^3 = (i-1, j+1),$$

$$\boxplus^4 = (i, j-1), \qquad\qquad \boxplus^5 = (i, j), \qquad\qquad \boxplus^6 = (i, j+1),$$

$$\boxplus^7 = (i+1, j-1), \qquad\qquad \boxplus^8 = (i+1, j), \qquad\qquad \boxplus^9 = (i+1, j+1).$$

On a quadrilateral grid, the second jet bundle of Y_d can be identified with

$$J^2 Y_d = X^\boxplus \times F^9, \tag{44}$$

with coordinates given by

$$\left\{ (\boxplus, y^a_{\boxplus l}) \,\middle|\, \boxplus \in X^\boxplus, \ y^a_{\boxplus l} \in \mathbb{R}, \ 1 \le l \le 9, \ 1 \le a \le \dim F \right\}, \tag{45}$$

in analogy with (39). The second jet prolongation of a discrete field φ_d is a map $j^2 \varphi_d : X^\boxplus \to J^2 Y_d$, defined as

$$j^2 \varphi_d(\boxplus) = \left(\boxplus, \varphi_{\boxplus^1}, \varphi_{\boxplus^2}, \varphi_{\boxplus^3}, \varphi_{\boxplus^4}, \varphi_{\boxplus^5}, \varphi_{\boxplus^6}, \varphi_{\boxplus^7}, \varphi_{\boxplus^8}, \varphi_{\boxplus^9} \right). \tag{46}$$

Similar to the continuous case, $J^2 Y_d$ can also be defined via iteration as the first jet bundle of $J^1 Y_d$. If we write coordinates on $J^1(J^1 Y_d)$ as

$$\left\{ (\boxplus, \square_m, y^a_{\square^l_m}) \,\middle|\, \boxplus \in X^\boxplus, \ \square_m \in \boxplus, \ y^a_{\square^l_m} \in \mathbb{R}, \ 1 \le l, m \le 4, \ 1 \le a \le \dim F \right\}, \tag{47}$$

then $J^2 Y_d$ consists of all elements of $J^1(J^1 Y_d)$, which satisfy

$$y^a_{\square^3_1} = y^a_{\square^4_2} = y^a_{\square^1_3} = y^a_{\square^2_4}, \qquad y^a_{\square^2_1} = y^a_{\square^1_2}, \qquad y^a_{\square^3_2} = y^a_{\square^2_3}, \qquad y^a_{\square^4_3} = y^a_{\square^3_4}, \qquad y^a_{\square^1_4} = y^a_{\square^4_1}, \tag{48}$$

for all a, which in our construction is always guaranteed due to (40) and (43). Therefore, $J^2 Y_d$ can be identified with $J^1(J^1 Y_d)$, unlike the continuous case, where $J^2 Y$ is strictly embedded into $J^1(J^1 Y)$.

2.5. Discrete action principle

The discretisation of the Lagrangian is based on the observation that the continuous action functional can be written as

$$\mathcal{A}[\varphi] = \sum_{\square \in X^\square} \int_{\text{Vol}(\square)} L(j^1 \varphi) \, d^{n+1}x, \tag{49}$$

where $\text{Vol}(\square) \subset X$ is the physical domain enclosed by $\square$. The integral in (49) is approximated by a function of values of φ_d in four different points in the spacetime grid, which corresponds to the discrete jet prolongation defined in (42). This is the discrete Lagrangian L_d, i.e.,

$$\int_{\text{Vol}(\square)} L(j^1 \varphi) \, d^{n+1}x \approx L_d(j^1 \varphi_d(\square)) = L_d(\square, \varphi_{\square^1}, \varphi_{\square^2}, \varphi_{\square^3}, \varphi_{\square^4}). \tag{50}$$

The action functional (49) is then approximated by

$$\mathcal{A}_d[\varphi_d] = \sum_{\square \in X^\square} L_d(j^1 \varphi_d(\square)), \tag{51}$$

which can also be written explicitly as

$$\mathcal{A}_d[\varphi_d] = \sum_{i=1}^{N_0-1} \sum_{j=1}^{N_1-1} L_d(\varphi_{i,j}, \varphi_{i,j+1}, \varphi_{i+1,j+1}, \varphi_{i+1,j}), \tag{52}$$

but for most of our derivations the abstract notation is more practical.

Specifically, the discrete Lagrangian is obtained by introducing a quadrature rule (e.g., trapezoidal, midpoint, Simpson) to approximate the integral in (50) as well as approximations of the fields and their derivatives. In the spirit of the Veselov discretisation [34–36,6], we will use the midpoint rule and first-order finite differences. This entails that the continuous fields in the Lagrangian are replaced with averages of field values at the four vertices of the grid cell $\square$, i.e.,

$$y^a \to \frac{1}{4}\left(y^a_{\square^1} + y^a_{\square^2} + y^a_{\square^3} + y^a_{\square^4}\right) \equiv \bar{y}^a(\square). \tag{53}$$

There are two possibilities for the definition of each of the derivatives. With reference to Fig. 2(b), let the vertical dimension correspond to x^0 (time) and the horizontal dimension to x^1 (space). Then ∂_0 can be defined along the left as well as along the right edge of the grid cell. Similarly, ∂_1 can be defined along the upper as well as along the lower edge. Best results are usually obtained for the most symmetric discretisation of the Lagrangian. Therefore we use the average of the respective options and replace the derivatives in the Lagrangian according to

$$z^a_0 \to \frac{1}{2}\left(\frac{y^a_{\square^4} - y^a_{\square^1}}{h} + \frac{y^a_{\square^3} - y^a_{\square^2}}{h}\right) \equiv \bar{z}^a_0(\square), \tag{54}$$

$$z^a_1 \to \frac{1}{2}\left(\frac{y^a_{\square^2} - y^a_{\square^1}}{h} + \frac{y^a_{\square^3} - y^a_{\square^4}}{h}\right) \equiv \bar{z}^a_1(\square). \tag{55}$$

At last, in order to obtain the discrete equations of motion, we only need to apply a discrete version of Hamilton's principle of stationary action. The vertical transformation σ^ϵ is discretised in the same way as the fields φ, by considering its values over points (i, j) of the discrete base space X_d, i.e., we replace $\sigma^\epsilon(x, y)$ with $\sigma^\epsilon_d = (\sigma^\epsilon_{i,j}(y_{i,j}))$. The generating vector field is computed as in (3), but on grid points only,

$$\eta^a_{i,j} = \left.\frac{d(\sigma^\epsilon_{i,j})^a}{d\epsilon}\right|_{\epsilon=0}. \tag{56}$$

The discrete version of the action principle (6) then reads

$$\left.\frac{d}{d\epsilon}\mathcal{A}_d[\sigma^\epsilon_d \circ \varphi_d]\right|_{\epsilon=0} = 0 \quad \text{for every } \sigma^\epsilon_d. \tag{57}$$

The explicit computation of (57) leads to

$$\left.\frac{d}{d\epsilon}\mathcal{A}_d[\sigma^\epsilon_d \circ \varphi_d]\right|_{\epsilon=0} = \sum_{\square \in X^\square} \left.\frac{d}{d\epsilon} L_d(j^1 \varphi^\epsilon_d(\square))\right|_{\epsilon=0}$$

$$= \sum_{\square \in X^\square} \sum_{l=1}^{4} \frac{\partial L_d}{\partial y^a_{\square^l}}(j^1 \varphi_d(\square)) \cdot \eta^a_{\square^l}(\varphi_{\square^l}). \tag{58}$$

As the variation of the action has to vanish for each $\eta^a_{i,j}$ on the spacetime grid independently, it is sufficient to consider only those contributions that are multiplied with the vector field at a fixed grid point (i, j). In total there are four such contributions (see Fig. 3),
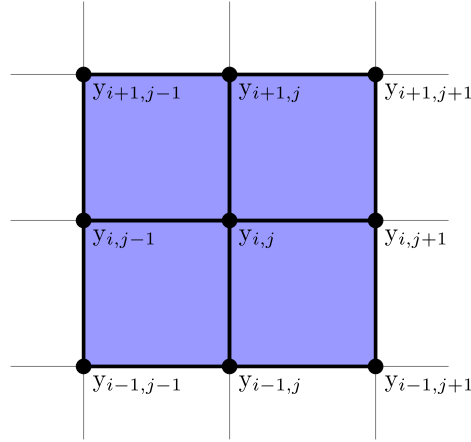
Fig. 3. Contributions to the discrete Euler–Lagrange field equations. The derivatives of the discrete Lagrangian at the highlighted grid cells with respect to $y_{i,j}$ correspond to the terms in Eq. (59).

$$\frac{d}{d\epsilon} \mathcal{A}_d[\sigma_d^\epsilon \circ \varphi_d]\bigg|_{\epsilon=0} =$$

$$\cdots + \frac{\partial L_d}{\partial y^a_{\square^1}} \Big(\varphi_{i,j}, \varphi_{i,j+1}, \varphi_{i+1,j+1}, \varphi_{i+1,j}\Big) \cdot \eta^a_{i,j}(\varphi_{i,j}) + \cdots$$

$$\cdots + \frac{\partial L_d}{\partial y^a_{\square^2}} \Big(\varphi_{i,j-1}, \varphi_{i,j}, \varphi_{i+1,j}, \varphi_{i+1,j-1}\Big) \cdot \eta^a_{i,j}(\varphi_{i,j}) + \cdots$$

$$\cdots + \frac{\partial L_d}{\partial y^a_{\square^3}} \Big(\varphi_{i-1,j-1}, \varphi_{i-1,j}, \varphi_{i,j}, \varphi_{i,j-1}\Big) \cdot \eta^a_{i,j}(\varphi_{i,j}) + \cdots$$

$$\cdots + \frac{\partial L_d}{\partial y^a_{\square^4}} \Big(\varphi_{i-1,j}, \varphi_{i-1,j+1}, \varphi_{i,j+1}, \varphi_{i,j}\Big) \cdot \eta^a_{i,j}(\varphi_{i,j}) + \cdots = 0.$$

The variation of the discrete action vanishes, if the sum of all expressions multiplying $\eta^a_{i,j}$ vanishes identically for all (i, j) (since the $(\sigma^\epsilon_{i,j})^a$ and therefore the $\eta^a_{i,j}$ are arbitrary). For each a, this requirement yields the discrete Euler–Lagrange field equations at (i, j),

$$0 = \frac{\partial L_d}{\partial y^a_{\square^1}} \Big(\varphi_{i,j}, \varphi_{i,j+1}, \varphi_{i+1,j+1}, \varphi_{i+1,j}\Big) + \frac{\partial L_d}{\partial y^a_{\square^2}} \Big(\varphi_{i,j-1}, \varphi_{i,j}, \varphi_{i+1,j}, \varphi_{i+1,j-1}\Big)$$

$$+ \frac{\partial L_d}{\partial y^a_{\square^3}} \Big(\varphi_{i-1,j-1}, \varphi_{i-1,j}, \varphi_{i,j}, \varphi_{i,j-1}\Big) + \frac{\partial L_d}{\partial y^a_{\square^4}} \Big(\varphi_{i-1,j}, \varphi_{i-1,j+1}, \varphi_{i,j+1}, \varphi_{i,j}\Big), \tag{59}$$

which can be compactly written as an equation on the discrete second jet bundle $J^2 Y_d$, namely

$$\big(j^2 \varphi_d(\boxplus)\big)^* \big(D_{EL}(L_d)\big)^a = \sum_{\substack{l=1 \\ \square^l = \boxplus^5}}^{4} \frac{\partial L_d}{\partial y^a_{\square^l}} \big(j^1 \varphi_d(\square)\big) = 0 \quad \text{for all } a \text{ and all } \boxplus. \tag{60}$$

Here, $D_{EL}(L_d)$ is the discrete Euler–Lagrange operator acting on the discrete Lagrangian L_d, which is evaluated on the prolongation $j^2 \varphi_d(\boxplus)$ as indicated by the pull-back notation. These relations define the variational integrator for a first-order Lagrangian field theory according to the Veselov discretisation of the Lagrangian as it was described above.

2.6. Discrete Noether theorem

Following the derivation of the continuous theory from the previous section, we consider vertical transformations $\sigma^\epsilon_{i,j}(y_{i,j})$ and define

$$\varphi^\epsilon_{i,j} = \sigma^\epsilon_{i,j} \circ \varphi_{i,j} \quad \text{with } \sigma^\epsilon_{i,j}|_{\epsilon=0} = \text{id and } \eta^a_{i,j} = \frac{d(\sigma^\epsilon_{i,j})^a}{d\epsilon}\bigg|_{\epsilon=0}. \tag{61}$$

The transformation $\sigma^\epsilon_{i,j}$ is a symmetry for the discrete Lagrangian (50) if

$$L_d\big(j^1 \varphi^\epsilon_d(\square)\big) = L_d\big(j^1 \varphi_d(\square)\big), \tag{62}$$

which is equivalent to the infinitesimal symmetry condition

$$\frac{d}{d\epsilon} L_d\big(j^1 \varphi^\epsilon_d(\square)\big)\bigg|_{\epsilon=0} = \sum_{l=1}^{4} \frac{\partial L_d}{\partial y^a_{\square^l}} \big(j^1 \varphi_d(\square)\big) \cdot \eta^a_{\square^l}(\varphi_{\square^l}) = 0. \tag{63}$$
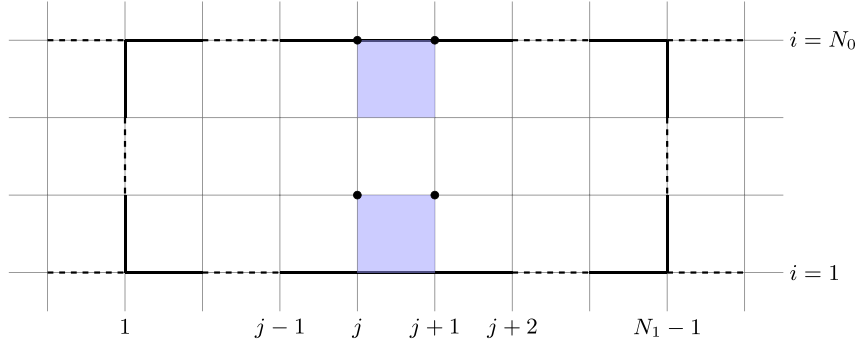
Fig. 4. Contributions to the discrete Noether theorem.

On each grid cell $\square$ we can define four discrete momentum maps in analogy to (20),

$$\left(\mathbf{j}^1 \varphi_d(\square)\right)^* J_{\square^l} = \frac{\partial L_d}{\partial \mathbf{y}^a_{\square^l}} \left(\mathbf{j}^1 \varphi_d(\square)\right) \cdot \eta^a_{\square^l}(\varphi_{\square^l}), \quad l \in \{1, 2, 3, 4\}. \tag{64}$$

Instead of two components of J, corresponding to the coordinates of the base space, we now have four contributions, corresponding to the vertices of a grid cell. With that, we can write the discrete symmetry condition (63) as

$$J_{\square^1} + J_{\square^2} + J_{\square^3} + J_{\square^4} = 0, \quad \text{for every } \square \in X^\square. \tag{65}$$

In addition to (64), we define some explicit shorthand notation that will become useful in the following derivation,

$$J^1_{i,j} = \frac{\partial L_d}{\partial \mathbf{y}^a_{\square^1}} \left(\varphi_{i,j}, \varphi_{i,j+1}, \varphi_{i+1,j+1}, \varphi_{i+1,j}\right) \cdot \eta^a_{i,j}(\varphi_{i,j}), \tag{66a}$$

$$J^2_{i,j} = \frac{\partial L_d}{\partial \mathbf{y}^a_{\square^2}} \left(\varphi_{i,j}, \varphi_{i,j+1}, \varphi_{i+1,j+1}, \varphi_{i+1,j}\right) \cdot \eta^a_{i,j+1}(\varphi_{i,j+1}), \tag{66b}$$

$$J^3_{i,j} = \frac{\partial L_d}{\partial \mathbf{y}^a_{\square^3}} \left(\varphi_{i,j}, \varphi_{i,j+1}, \varphi_{i+1,j+1}, \varphi_{i+1,j}\right) \cdot \eta^a_{i+1,j+1}(\varphi_{i+1,j+1}), \tag{66c}$$

$$J^4_{i,j} = \frac{\partial L_d}{\partial \mathbf{y}^a_{\square^4}} \left(\varphi_{i,j}, \varphi_{i,j+1}, \varphi_{i+1,j+1}, \varphi_{i+1,j}\right) \cdot \eta^a_{i+1,j}(\varphi_{i+1,j}). \tag{66d}$$

In order to derive the discrete conservation law, we follow the argument from Section 2.3. If the Lagrangian is invariant under a given vertical transformation, then the action is also invariant under this transformation, namely,

$$\frac{d}{d\epsilon} \mathcal{A}_d\left[\varphi^\epsilon_d\right]\bigg|_{\epsilon=0} = 0, \tag{67}$$

which becomes

$$\frac{d}{d\epsilon} \mathcal{A}_d\left[\varphi^\epsilon_d\right]\bigg|_{\epsilon=0} = \sum_{\square \in X^\square} \sum_{l=1}^{4} \frac{\partial L_d}{\partial \mathbf{y}^a_{\square^l}} \left(\mathbf{j}^1 \varphi_d(\square)\right) \cdot \eta^a_{\square^l}(\varphi_{\square^l})$$

$$= \sum_{\square \in X^\square} \sum_{l=1}^{4} \left(\mathbf{j}^1 \varphi_d(\square)\right)^* J_{\square^l} = 0. \tag{68}$$

Those contributions to the sum that originate from interior points vanish in virtue of the discrete Euler–Lagrange field equations (59). Specifically,

$$J^1_{i,j} + J^2_{i,j-1} + J^3_{i-1,j-1} + J^4_{i-1,j} = \left(\mathbf{j}^2 \varphi_d(\boxplus)\right)^* \left(D_{\mathrm{EL}}(L_d)\right)^a_{i,j} \cdot \eta^a_{i,j}(\varphi_{i,j}) = 0, \tag{69}$$

so that only boundary cells contribute to (68), c.f. Fig. 4. For simplicity, we will assume periodic boundary conditions in the space-like variable, then the only contribution comes from the boundary of the time-like variable,

$$0 = \sum_{j=1}^{N_1-1} \left[J^1_{1,j} + J^2_{1,j-1} + J^3_{N_0-1,j-1} + J^4_{N_0-1,j}\right]. \tag{70}$$

Using the discrete symmetry condition (65) to replace the first two terms as well as the periodicity of the spatial domain, we obtain the discrete counterpart of (24), that is the discrete conservation law

$$\sum_{j=1}^{N_1-1} \left[J^3_{1,j} + J^4_{1,j}\right] = \sum_{j=1}^{N_1-1} \left[J^3_{N_0-1,j} + J^4_{N_0-1,j}\right], \tag{71}$$

see Fig. 4. Since the number of timesteps N_0 is arbitrary, this is equivalent to

$$\mathcal{J}_d = \sum_{\substack{j=1 \\ \square^1=(i,j)}}^{N_1-1} \left[J_{\square^3}\left(\mathrm{j}^1 \varphi_d(\square)\right) + J_{\square^4}\left(\mathrm{j}^1 \varphi_d(\square)\right) \right] = \text{const.} \quad \text{for all } i, \tag{72}$$

in analogy to (25). This proves that variational integrators preserve discrete invariants $\mathcal{J}_d$ to machine accuracy.[3]

Discrete divergence symmetries

We can generalise the discrete Noether theorem in a similar way as we did with the continuous Noether theorem for divergence symmetries. In analogy to (26), we rewrite the discrete symmetry condition (65) as

$$J_{\square^1} + J_{\square^2} + J_{\square^3} + J_{\square^4} = B_{\square^1} + B_{\square^2} + B_{\square^3} + B_{\square^4}. \tag{73}$$

Making no assumption on the $B_{\square^l}$, Eq. (70) takes the form

$$\sum_{j=1}^{N_1-1} \left[J_{1,j}^1 + J_{1,j-1}^2 - B_{1,j}^1 - B_{1,j-1}^2 + J_{N_0-1,j-1}^3 + J_{N_0-1,j}^4 - B_{N_0-1,j-1}^3 - B_{N_0-1,j}^4 \right]$$

$$= \sum_{i=2}^{N_0-1} \sum_{j=1}^{N_1-1} \left[B_{i,j}^1 + B_{i,j-1}^2 + B_{i-1,j-1}^3 + B_{i-1,j}^4 \right]. \tag{74}$$

We conclude, that we have the discrete equivalent of a divergence symmetry, if the sum of all $B_{i,j}^l$ in the interior of X_d vanishes,

$$\sum_{i=2}^{N_0-1} \sum_{j=1}^{N_1-1} \left[B_{i,j}^1 + B_{i,j-1}^2 + B_{i-1,j-1}^3 + B_{i-1,j}^4 \right] = 0, \tag{75}$$

such that only the fluxes at the boundaries of X_d contribute to (74). Then, the generalised Noether charge (72) becomes

$$\mathcal{J}_d = \sum_{\substack{j=1 \\ \square^1=(i,j)}}^{N_1-1} \left[J_{\square^3}\left(\mathrm{j}^1 \varphi_d(\square)\right) + J_{\square^4}\left(\mathrm{j}^1 \varphi_d(\square)\right) - B_{\square^3}\left(\mathrm{j}^1 \varphi_d(\square)\right) - B_{\square^4}\left(\mathrm{j}^1 \varphi_d(\square)\right) \right] = \text{const.} \quad \text{for all } i, \tag{76}$$

in correspondence with (32).

3. Formal Lagrangians

In this section we review the idea of formal (or adjoint) Lagrangians [18,19] and the corresponding Noether theorem as introduced by Ibragimov [20]. We provide a geometric view of this theory in line with the formalism adopted in the rest of the paper, so that the application of the results of Section 2 to formal Lagrangians becomes natural. At last, we propose a discrete version of Ibragimov's extension of Noether's theorem, which allows us to prove discrete conservation laws for variational integrators obtained from formal Lagrangians.

3.1. Inverse problem of the calculus of variations

Consider a generic system of nonlinear partial differential equations,

$$\mathcal{F}(\mathrm{x}, u, u_{(1)}, \ldots, u_{(k)}) = 0, \tag{77}$$

for a field $u : X \to F$, where $u_{(k)}$ denotes all derivatives of order k. Here, $\mathcal{F}$ is a (sufficiently regular) function of the point $\mathrm{x} \in X$, the field $u(\mathrm{x})$, and its derivatives up to order k, taking values in $\mathbb{R}^m$, where $m = \dim F$, so that we have as many equations as variables. The components of $\mathcal{F}$ are denoted $\mathcal{F}_a$ with $1 \le a \le m$.

Let B be the space of functions $u : X \to F$ where the solution of (77) is sought. We assume that B is a Banach space and denote by B^* its topological dual, i.e., the space of linear continuous functionals from B to $\mathbb{R}$. The specific choice of the space B depends on the problem at hand. For definiteness, we require $B \hookrightarrow [L^2(X)]^m$, i.e., B is continuously embedded in $[L^2(X)]^m$. If the function $\mathrm{x} \mapsto \mathcal{F}(\mathrm{x}, u, u_{(1)}, \ldots, u_{(k)})$ defined in (77) belongs to $[L^2(X)]^m$, then it can be identified with the mapping $\mathcal{F} : B \to B^*$ defined by

$$\langle \mathcal{F}[u], v \rangle = \int_X v(\mathrm{x}) \cdot \mathcal{F}(\mathrm{x}, u, u_{(1)}, \ldots, u_{(k)}) \, d^{n+1}\mathrm{x}, \tag{78}$$

where $\langle \cdot, \cdot \rangle : B^* \times B \to \mathbb{R}$ is the duality pairing between elements of B^* and B. With some abuse of notation we denote by $\mathcal{F}$ both the function in (77), which is defined on a finite-dimensional space, and the mapping $\mathcal{F} : B \to B^*$, which is defined on a Banach space.

The nonlinear operator $\mathcal{F}$ admits a natural variational formulation if there exists a functional $\mathcal{A} : B \to \mathbb{R}$ of class $C^1(B)$, such that $\mathcal{F} = D\mathcal{A}$, where D denotes the Fréchet derivative operator. We recall that the Fréchet derivative of $\mathcal{A} : B \to \mathbb{R}$ is a map $D\mathcal{A} : B \to B^*$,

---

[3] In practice, many variational integrators require the solution of nonlinear algebraic equations. In that case, the accuracy to which the invariants are preserved will generally depend on the accuracy of the nonlinear solver.

such that for every $u, v \in B$, $\langle D\mathcal{A}[u], v \rangle$ coincides with the Gateaux derivative of $\mathcal{A}$ along v [33,37], that is

$$\langle D\mathcal{A}[u], v \rangle = \lim_{\epsilon \to 0} \frac{\mathcal{A}[u + \epsilon v] - \mathcal{A}[u]}{\epsilon} = \frac{d}{d\epsilon} \mathcal{A}[u + \epsilon v] \Big|_{\epsilon=0}. \tag{79}$$

In this framework, the functional derivative $\delta\mathcal{A}[u]/\delta u$, if it exists, is defined as the L^2-representative of the Fréchet derivative. Specifically, $\delta\mathcal{A}[u]/\delta u$ is the unique element of $\left[L^2(X)\right]^m$, if it exists, such that

$$\langle D\mathcal{A}[u], v \rangle = \int_X \frac{\delta\mathcal{A}[u]}{\delta u} \cdot v \, dx, \tag{80}$$

for all $v \in B$ [38]. Here, we are implicitly assuming that boundary conditions on the field u are encoded in the definition of the linear space B; this is possible in particular for both homogeneous Dirichlet and periodic boundary conditions.

Analogously to (79), the Fréchet derivative at $u \in B$ of a functional $\mathcal{F} : B \to B^*$ of class C^1 is the linear continuous operator $D\mathcal{F}[u] : B \to B^*$, such that for every $v \in B$, $D\mathcal{F}[u]v$ coincides with the Gateaux derivative of $\mathcal{F}$ along v, i.e.,

$$D\mathcal{F}[u]v = \frac{d}{d\epsilon} \mathcal{F}[u + \epsilon v] \Big|_{\epsilon=0}. \tag{81}$$

Therefore, $D\mathcal{F} : B \to \mathcal{L}(B, B^*) \cong B^* \otimes B^*$, where $\mathcal{L}(B, B^*)$ is the space of linear continuous operators from B to B^*. If $\mathcal{F}$ admits a variational formulation, then the second-order Fréchet derivative $D\mathcal{F} = D^2\mathcal{A} : B \to B^* \otimes B^*$ should be symmetric at each u, hence a necessary condition for the existence of the action $\mathcal{A}$ is

$$\langle D\mathcal{F}[u]v, w \rangle = \langle D\mathcal{F}[u]w, v \rangle \quad \forall u, v, w \in B. \tag{82}$$

It turns out that this condition is also sufficient [33,39,40].

Unfortunately, for many interesting systems this condition is not fulfilled. Some of these systems admit a variational principle after a variable transformation. This is the case for many equations from fluid dynamics [41]. But such transformations can be inconvenient since the new variables might suffer from problems with respect to non-uniqueness, boundary conditions or regularity. In addition, the resulting variational principle might be subject to constraints on the variations which are not easily dealt with at the discrete level. Instead, they call for extensions of the theory that complicate its application [42,43]. For a still larger class of systems, a variational formulation is not known, even after a change of coordinates. Nevertheless, we can derive variational integrators for these systems by considering the following construction.

3.2. Formal Lagrangians

A generic system of differential equations (77) can be treated as part of a Lagrangian system, constructed by doubling the number of variables. The action of such a Lagrangian system is

$$\mathcal{A}[u, v] = \langle \mathcal{F}[u], v \rangle, \tag{83}$$

with $u, v \in B$ and the pairing as defined in (78). Then, the Lagrangian is

$$L(x, u, u_{(1)}, \ldots, u_{(k)}, v, v_{(1)}, \ldots, v_{(k)}) = v(x) \cdot \mathcal{F}(x, u, u_{(1)}, \ldots, u_{(k)}). \tag{84}$$

This is referred to as formal (or adjoint) Lagrangian [18,19]. The variational principle (6) applied to (84) gives the original equation,

$$\frac{\delta\mathcal{A}[u, v]}{\delta v} = \mathcal{F}[u] = 0, \tag{85}$$

as well as the *adjoint equation*, which is defined by

$$\frac{\delta\mathcal{A}[u, v]}{\delta u} = \mathcal{F}^*[u, v] = 0, \tag{86}$$

assuming that the functional derivative exists.

Summarising, if we consider a dynamical system described by the variables (u, v), whose dynamics are governed by the equations $\mathcal{F}[u] = 0$ and $\mathcal{F}^*[u, v] = 0$, then this system has a variational formulation and contains the original system (77) as a subsystem. In the following, we refer to the system (85)–(86) as *extended system* describing the dynamics of the variables (u, v).

Geometry of formal Lagrangians

As before, we identify a field u with a section $\varphi : X \to Y$ of the configuration bundle Y. With $j^k\varphi$ the kth jet prolongation of φ, we can write the system of equation (77) as a function on the jet bundle J^kY,

$$\mathcal{F}(j^k\varphi) = 0. \tag{87}$$

The configuration bundle $\tilde{Y}$ of the extended system is given by the Cartesian product

$$\tilde{Y} = X \times F \times F, \tag{88}$$

together with the projection

$$\tilde{\pi} : \tilde{Y} \to X. \tag{89}$$

Local coordinates on $\tilde{Y}$ are denoted by $(x^\mu, y^a, \tilde{y}^a)$ with $1 \leq a \leq m = \dim F$. Sections $\tilde{\varphi} : X \to \tilde{Y}$ are at the same time associated to the graph of u and v,

$$\tilde{\varphi} : x \mapsto \left(x^\mu, \varphi^a(x), \tilde{\varphi}^a(x)\right) = \left(x^\mu, u^a(x), v^a(x)\right), \tag{90}$$

where the φ^a are associated with the physical variables u^a and the $\tilde{\varphi}^a$ are associated with the adjoint variables v^a. With $j^k\tilde{\varphi}$ the kth jet prolongation of $\tilde{\varphi}$, we can write the formal Lagrangian (84) as a function on the jet bundle $J^k\tilde{Y}$,

$$L\left(j^k\tilde{\varphi}\right) = v \cdot \mathcal{F}\left(x, u, u_{(1)}, \ldots, u_{(k)}\right), \tag{91}$$

and the formal action functional (83) as

$$\mathcal{A}[\tilde{\varphi}] = \int_X L\left(j^k\tilde{\varphi}\right) d^{n+1}x. \tag{92}$$

In the following we shall consider first-order Lagrangians only, i.e., $k = 1$.

3.3. Noether theorem

Ibragimov [20] has shown that the adjoint equations (86) inherit all symmetries of the original equations (85). Therefore it is possible to determine conservation laws of any system of differential equations, even without a natural Lagrangian, by application of the Noether theorem from Section 2.3 to the corresponding formal Lagrangian (91). Again, we restrict our attention to the case of first order field theories.

More specifically, if the original system (77) admits a symmetry related to the infinitesimal vector field

$$V = \eta^a(x, y)\frac{\partial}{\partial y^a}, \tag{93}$$

then the extended system (85)–(86) admits a symmetry related to the vector field

$$\tilde{V} = \eta^a(x, y)\frac{\partial}{\partial y^a} + \tilde{\eta}^a(x, y, \tilde{y})\frac{\partial}{\partial \tilde{y}^a}, \tag{94}$$

with appropriately chosen coefficients $\tilde{\eta}^a$ [20]. By definition, the vector field (93) describes a Lie point symmetry of (77), if there exists a matrix-valued function $\lambda = (\lambda_a^b(x, y, z))$, such that

$$j^1V(\mathcal{F}_a) = \lambda_a^b \mathcal{F}_b, \tag{95}$$

where $\mathcal{F}(x, y, z)$ is the function defining the system of first order equations $\mathcal{F}(x, u, u_{(1)})$. We apply the jet prolongation of the extended vector field (94) to the Lagrangian (91) of the extended system (85)–(86), treating $\tilde{\eta}^a$ as unknown coefficients,

$$j^1\tilde{V}(L) = \tilde{\eta}^a \mathcal{F}_a + \tilde{y}^a j^1V(\mathcal{F}_a) = \tilde{\eta}^a \mathcal{F}_a + \tilde{y}^a \lambda_a^b \mathcal{F}_b. \tag{96}$$

We want to choose the $\tilde{\eta}^a$ so that the symmetry condition (17) is satisfied. We see that it is sufficient to set

$$\tilde{\eta}^a = -\lambda_b^a \tilde{y}^b \tag{97}$$

for (96) to vanish. The extended vector field (94) therefore becomes

$$\tilde{V} = \eta^a \frac{\partial}{\partial y^a} - \lambda_b^a \tilde{y}^b \frac{\partial}{\partial \tilde{y}^a}, \tag{98}$$

with λ given in the symmetry condition (95). In the line of the standard Noether theorem, c.f. Section 2.3, if (94) is a symmetry of the formal Lagrangian (91), we obtain the local conservation law

$$\operatorname{div}\tilde{J}(j^1\tilde{\varphi}) = 0, \tag{99}$$

for all solutions $\tilde{\varphi}$ of the Euler–Lagrange equations on $\tilde{Y}$, with the Noether current $\tilde{J}$ given by

$$\tilde{J}^\mu\left(j^1\tilde{\varphi}\right) = \frac{\partial L}{\partial z_\mu^a}\left(j^1\tilde{\varphi}\right) \cdot \eta^a(\tilde{\varphi}) + \frac{\partial L}{\partial \tilde{z}_\mu^a}\left(j^1\tilde{\varphi}\right) \cdot \tilde{\eta}^a(\tilde{\varphi}). \tag{100}$$

For formal Lagrangians of the form (91), the second terms on the right-hand side are zero. Sometimes, however, we might want to symmetrise the Lagrangian (see e.g. Section 4.3), in which case these terms do contribute to the Noether current. From Eq. (100), we obtain the (global) Noether charge (25),

$$\tilde{\mathcal{J}} = \int \left[\frac{\partial L}{\partial z_t^a}\left(j^1\tilde{\varphi}\right) \cdot \eta^a(\tilde{\varphi}) + \frac{\partial L}{\partial \tilde{z}_t^a}\left(j^1\tilde{\varphi}\right) \cdot \tilde{\eta}^a(\tilde{\varphi})\right] d^n x, \tag{101}$$

which is constant for all times t. Usually, formal Lagrangians of the form (91) do not feature a time derivative of the adjoint variables, so that the Noether charge simplifies accordingly.

Restriction of conservation laws

The conservation law thus obtained is of course a conservation law of the extended system (85)–(86), and therefore generally depends on both u and v. In order to restrict conservation laws of the extended systems to conservation laws of the physical system, we have to

find a suitable way of restricting solutions of the extended system to solutions of the physical system. For that purpose, the following two definitions will become useful.

If the adjoint equation (86), restricted to $v = u$, becomes equivalent to the original equation (77), i.e.,

$$\mathcal{F}^*[u, u] = \lambda \mathcal{F}[u], \tag{102}$$

for some matrix λ, possibly depending on the fields and their derivatives, the system (77) is called self-adjoint in the sense of Ibragimov [19]. If $\mathcal{F}[u]$ is a linear operator and λ the identity matrix $\mathbb{1}_{m \times m}$, the above definition coincides with the standard definition of formally self-adjoint operators. In general, however, the systems we are considering are not self-adjoint. Ibragimov relaxed the requirement of self-adjointness by introducing the concept of quasi-self-adjointness [21,44]. This is a generalisation of self-adjointness where $v = u$ is generalised to $v = \phi(u)$ for some function $\phi : F \to F$. The advantage of self-adjointness or quasi-self-adjointness in the sense of Ibragimov is the possibility to build a solution of the full extended system (85)–(86) given a solution of the original problem (85).

Given a diffeomorphism $\phi : F \to F$ of F into itself, we can build the embedding

$$\begin{aligned} \Phi \quad : \quad & Y \quad \hookrightarrow \quad \tilde{Y}, \\ & (x, y) \quad \mapsto \quad (x, y, \phi(y)). \end{aligned} \tag{103}$$

In order to use this embedding to restrict the Noether current to the physical system, we need to lift Φ to a map from $J^1 Y$ to $J^1 \tilde{Y}$. With this aim, let us consider a generic section φ of Y for which the composed map

$$\tilde{\varphi} = \Phi \circ \varphi : X \to \tilde{Y} \tag{104}$$

amounts to

$$\tilde{\varphi} : x \mapsto \big(x, u(x), \phi(u(x))\big) = \big(x, u(x), v(x)\big), \tag{105}$$

where $u(x)$ is the field component corresponding to φ, and we have defined the second field $v = \phi \circ u : X \to F$. It follows that the condition $\tilde{\varphi} \circ \tilde{\pi} = \mathrm{id}_X$ is satisfied and $\tilde{\varphi} = \Phi \circ \varphi$ is indeed a section of $\tilde{Y}$, i.e., the composition with Φ maps sections of Y into sections of $\tilde{Y}$. By the chain rule, we compute the first jet of $\tilde{\varphi}$, that is

$$j^1 \tilde{\varphi}(x) = \big(x, u(x), v(x), Du(x), Dv(x)\big)$$

$$= \big(x, u(x), \phi(u(x)), Du(x), D\phi(u(x)) \cdot Du(x)\big), \tag{106}$$

where Df denotes the Jacobian matrix of a function f. We now can define the lift

$$j^1 \Phi : J^1 Y \to J^1 \tilde{Y} \tag{107}$$

by

$$j^1 \Phi (x, y, z) = \big(x, y, \phi(y), z, D\phi(y) \cdot z\big), \tag{108}$$

which by construction satisfies the identity

$$j^1 (\Phi \circ \varphi) = j^1 \Phi \circ j^1 \varphi, \tag{109}$$

as desired. With an abuse of notation, the symbol $j^1 \Phi$ is not used here in the usual sense of jet prolongation, instead $j^1 \Phi$ is a lift of Φ up to the first jet bundle. We can now use (108) to pull back the Noether current in (99) if we assume that $\tilde{\varphi}$ can be realised in the form $\tilde{\varphi} = \Phi \circ \varphi$,

$$\tilde{J}(j^1 \tilde{\varphi}) = \tilde{J}\big(j^1 (\Phi \circ \varphi)\big) = \tilde{J}\big(j^1 \Phi \circ j^1 \varphi\big) = \big(\tilde{J} \circ j^1 \Phi\big)\big(j^1 \varphi\big). \tag{110}$$

If $\tilde{\varphi} = \Phi \circ \varphi$ solves the equations of the extended system (85)–(86), then $\mathrm{div}\, \tilde{J}(j^1 \tilde{\varphi}) = 0$. Upon defining the restricted Noether current by

$$J = \tilde{J} \circ j^1 \Phi, \tag{111}$$

the Noether theorem (99) takes the form

$$\mathrm{div}\, J(j^1 \varphi) = 0, \tag{112}$$

which expresses the local conservation law for the physical system.

This result is crucial for the application of variational integrators to formal Lagrangians. Without the construction of a solution $(u, \phi(u))$ of the extended system from a solution u of the physical system, it is in general not possible to determine the discrete momenta that are conserved by the variational integrator due to symmetries of the physical system.

Generalisations

To simplify the derivations and the analysis of the conservation laws, it is sometimes useful to add a term $G(j^k \varphi)$ to the Lagrangian, that is a function of the coordinates, the physical variables and their derivatives, but not of the adjoint variables and their derivatives. The modified Lagrangian,

$$L'(j^k \tilde{\varphi}) = L(j^k \tilde{\varphi}) + G(j^k \tilde{\varphi}) \quad \text{with } \frac{\partial G}{\partial \tilde{y}^a}(j^k \tilde{\varphi}) = 0 \text{ and } \frac{\partial G}{\partial \tilde{z}^a_\mu}(j^k \tilde{\varphi}) = 0 \quad \text{for all } \mu, a, \tag{113}$$

yields the same physical equations of motion as L, but in general will lead to different adjoint equations. This freedom can be used to simplify the search for the embedding Φ used to restrict the conservation laws of the extended system to the physical system.

For example, for equations of advection type,

$$u_t + A(u, u_x) = 0, \tag{114}$$

where A is some possibly nonlinear function of the field u and its spatial derivatives, it is possible to construct an extended system which is guaranteed to be always self-adjoint. It suffices to choose $G = -u \cdot A$, so that the Lagrangian becomes

$$L'\big(j^k \tilde{\varphi}\big) = v \cdot u_t + v \cdot A(u, u_x) - u \cdot A(u, u_x). \tag{115}$$

While this simplifies the construction of the embedding Φ, it complicates the extension of symmetries. In general it will not be possible to determine the components of the generating vector field in an algorithmic way as in (97).

Discrete embedding

At this point, the discretisation of formal Lagrangians can be carried out straightforwardly as an application of the theory reviewed in Section 2. The only missing element is the restriction of discrete conservation laws of the extended system to the physical system.

Since the fibres of Y_d are copies of F, we can consider a diffeomorphism $\phi : F \to F$, as in the continuous case and define the discrete embedding

$$\begin{array}{rlll} \Phi_d \; : & Y_d & \hookrightarrow & \tilde{Y}_d, \\ & (y_{i,j}) & \mapsto & \big(y_{i,j}, \, \phi(y_{i,j})\big), \end{array} \tag{116}$$

and its lift to the discrete first jet bundle,

$$j^1 \Phi_d : J^1 Y_d \to J^1 \tilde{Y}_d, \tag{117}$$

which is given by

$$j^1 \Phi_d(\square) = \big(\square, \; y_{\square^1}, y_{\square^2}, y_{\square^3}, y_{\square^4}, \; \phi(y_{\square^1}), \phi(y_{\square^2}), \phi(y_{\square^3}), \phi(y_{\square^4})\big). \tag{118}$$

With this we can pullback discrete conservation laws (72) of the extended system,

$$\tilde{\mathcal{J}}_d = \sum_{\substack{j=1 \\ \square^1 = (i,j)}}^{N_1 - 1} \left[\tilde{J}_{\square^3}\big(j^1 \tilde{\varphi}_d(\square)\big) + \tilde{J}_{\square^4}\big(j^1 \tilde{\varphi}_d(\square)\big) \right] = \text{const.} \quad \text{for all } i, \tag{119}$$

to the physical system by defining the restricted discrete momentum map

$$J_{\square^l} = \tilde{J}_{\square^l} \circ j^1 \Phi_d(\square), \tag{120}$$

such that the restricted Noether charge becomes

$$\mathcal{J}_d = \sum_{\substack{j=1 \\ \square^1 = (i,j)}}^{N_1 - 1} \left[J_{\square^3}\big(j^1 \varphi_d(\square)\big) + J_{\square^4}\big(j^1 \varphi_d(\square)\big) \right] = \text{const.} \quad \text{for all } i. \tag{121}$$

It is worth noticing that we do not pull back the Noether charge $\tilde{\mathcal{J}}$ but the momentum maps $\tilde{J}_{\square^l}$, just as in the continuous case (111), where we pull back the Noether current $\tilde{J}$.

4. Applications

In this section we present two applications, the linear advection equation and the vorticity equation. The linear advection equation in one spatial dimension is a prototypical example that shares many characteristics with more complicated systems from plasma physics and fluid dynamics. The vorticity equation in two spatial dimensions is an important equation of fluid dynamics which is widely used for example in atmospheric sciences but also in plasma physics. It consists of an advection equation which is coupled to a Poisson equation.

On the continuous level, we construct the formal Lagrangian and compute variations of the corresponding action functional. We check the resulting equations of motion for self-adjointness or quasi-self-adjointness in order to define the restriction map ϕ. Then we check the equations for symmetries and compute the extended generating vector field. This vector field is applied to the Lagrangian in order to compute the corresponding conservation laws. After discretising the formal Lagrangian, we derive the actual variational integrator and check the discrete equations of motion for discrete self-adjointness or quasi-self-adjointness. Then we check the discrete Lagrangian for symmetries and compute the discrete conservation laws. We will discuss different discretisations of the Lagrangian, especially with respect to different quadrature rules, and their consequences for properties of the resulting variational integrator (e.g., explicit vs. implicit) and the corresponding conservation laws.

Discretisation of formal Lagrangians

Before we proceed, we want to give a short comment on the discretisation of formal Lagrangians. A straight forward application of the Veselov discretisation, i.e., using the midpoint quadrature rule for both time and space integration, as it was presented in Section 2.5 and as it is used in many works on variational integrators, does not appear to be suitable for formal Lagrangians due to their particular structure. As we will see in Section 4.2 on the linear advection equation, it results in numerical schemes, which are prone to unphysical

oscillations in space. Instead, when discretising the derivatives with first-order finite differences, the spatial integral can be approximated by the trapezoidal or Simpson rule. Another option to avoid this issue is to move to a slightly more abstract Galerkin framework and use e.g. Lagrange polynomials [45] or splines [46] as basis functions.

4.1. Some definitions

So far, we have only considered the midpoint rule (Veselov discretisation) for approximating the Lagrangian $L(u, u_t, u_x)$, namely,

$$L_d^{\text{mp}}\big(\mathrm{j}^1\varphi_d(\Box)\big) = h_t h_x \, L\left(\tfrac{1}{4}\big(\varphi_{\Box^1} + \varphi_{\Box^2} + \varphi_{\Box^3} + \varphi_{\Box^4}\big), \tfrac{1}{2}\Big(\tfrac{\varphi_{\Box^4}-\varphi_{\Box^1}}{h_t} + \tfrac{\varphi_{\Box^3}-\varphi_{\Box^2}}{h_t}\Big), \tfrac{1}{2}\Big(\tfrac{\varphi_{\Box^2}-\varphi_{\Box^1}}{h_x} + \tfrac{\varphi_{\Box^3}-\varphi_{\Box^4}}{h_x}\Big)\right). \tag{122}$$

As will be discussed below, for formal Lagrangians the midpoint rule does not appear to lead to stable variational integrators. Therefore we will also consider some alternative discretisations of the Lagrangian. First, the trapezoidal rule,

$$L_d^{\text{tr}}\big(\mathrm{j}^1\varphi_d(\Box)\big) = \frac{h_t h_x}{4}\left(L\Big(\varphi_{\Box^1}, \tfrac{\varphi_{\Box^4}-\varphi_{\Box^1}}{h_t}, \tfrac{\varphi_{\Box^2}-\varphi_{\Box^1}}{h_x}\Big) + L\Big(\varphi_{\Box^2}, \tfrac{\varphi_{\Box^3}-\varphi_{\Box^2}}{h_t}, \tfrac{\varphi_{\Box^2}-\varphi_{\Box^1}}{h_x}\Big)\right.$$

$$\left. + L\Big(\varphi_{\Box^3}, \tfrac{\varphi_{\Box^3}-\varphi_{\Box^2}}{h_t}, \tfrac{\varphi_{\Box^3}-\varphi_{\Box^4}}{h_x}\Big) + L\Big(\varphi_{\Box^4}, \tfrac{\varphi_{\Box^4}-\varphi_{\Box^1}}{h_t}, \tfrac{\varphi_{\Box^3}-\varphi_{\Box^4}}{h_x}\Big)\right), \tag{123}$$

which leads to the (explicit) Leapfrog scheme. Second, a combination of both, the midpoint rule for time integration and the trapezoidal rule for spatial integration,

$$L_d^{\text{mt}}\big(\mathrm{j}^1\varphi_d(\Box)\big) = \frac{h_t h_x}{2}L\left(\tfrac{\varphi_{\Box^1}+\varphi_{\Box^4}}{2}, \tfrac{\varphi_{\Box^4}-\varphi_{\Box^1}}{h_t}, \tfrac{1}{2}\Big(\tfrac{\varphi_{\Box^2}-\varphi_{\Box^1}}{h_x} + \tfrac{\varphi_{\Box^3}-\varphi_{\Box^4}}{h_x}\Big)\right) + \frac{h_t h_x}{2}L\left(\tfrac{\varphi_{\Box^2}+\varphi_{\Box^3}}{2}, \tfrac{\varphi_{\Box^3}-\varphi_{\Box^2}}{h_t}, \tfrac{1}{2}\Big(\tfrac{\varphi_{\Box^2}-\varphi_{\Box^1}}{h_x} + \tfrac{\varphi_{\Box^3}-\varphi_{\Box^4}}{h_x}\Big)\right). \tag{124}$$

This leads to a simplified implicit scheme with the same conservation properties as the midpoint scheme derived from (122) but without the aforementioned instability.

We want to define some shorthand notation similar to (53)–(55),

$$u \to \frac{1}{4}\big(u_{\Box^1} + u_{\Box^2} + u_{\Box^3} + u_{\Box^4}\big) \equiv \bar{u}(\Box) \tag{125}$$

$$\frac{\partial u}{\partial t} \to \frac{1}{2}\left(\frac{u_{\Box^4} - u_{\Box^1}}{h_t} + \frac{u_{\Box^3} - u_{\Box^2}}{h_t}\right) \equiv \bar{u}_t(\Box), \tag{126}$$

$$\frac{\partial u}{\partial x} \to \frac{1}{2}\left(\frac{u_{\Box^2} - u_{\Box^1}}{h_x} + \frac{u_{\Box^3} - u_{\Box^4}}{h_x}\right) \equiv \bar{u}_x(\Box). \tag{127}$$

For the product of two fields as well as the product of a field and a derivative, according to the mixed use of trapezoidal and midpoint rule, we obtain

$$vu \to \frac{1}{2}\left[\frac{v_{\Box^1} + v_{\Box^4}}{2}\frac{u_{\Box^1} + u_{\Box^4}}{2} + \frac{v_{\Box^2} + v_{\Box^3}}{2}\frac{u_{\Box^2} + u_{\Box^3}}{2}\right] \equiv \overline{vu}(\Box), \tag{128}$$

$$v\frac{\partial u}{\partial t} \to \frac{1}{2}\left[\frac{v_{\Box^1} + v_{\Box^4}}{2}\frac{u_{\Box^4} - u_{\Box^1}}{h_t} + \frac{v_{\Box^2} + v_{\Box^3}}{2}\frac{u_{\Box^3} - u_{\Box^2}}{h_t}\right] \equiv \overline{vu_t}(\Box), \tag{129}$$

$$v\frac{\partial u}{\partial x} \to \frac{1}{2}\left[\frac{v_{\Box^1} + v_{\Box^2}}{2}\frac{u_{\Box^2} - u_{\Box^1}}{h_x} + \frac{v_{\Box^3} + v_{\Box^4}}{2}\frac{u_{\Box^3} - u_{\Box^4}}{h_x}\right] \equiv \overline{vu_x}(\Box). \tag{130}$$

This will prove handy in writing the discrete Lagrangians.

4.2. Linear advection equation

The initial value problem for the linear advection equation is an instructive example for the derivation of numerical schemes. Its analytic solution is known and its structure is similar to the systems we will consider in subsequent publications [22–24]. It is given by

$$u_t + c u_x = 0, \tag{131}$$

with initial condition

$$u(0, x) = u_0(x), \tag{132}$$

and describes the advection of a field $u(t, x)$ with a constant velocity c that is given as a parameter. The analytic solution is

$$u(t, x) = u_0(x - ct). \tag{133}$$

Here, x denotes the spatial variable only, while the base space X comprises both space and time, i.e., $(t, x) \in X$. The field u can be interpreted as a density, such that $\int cu\,dx$ is the total momentum and $\frac{1}{2}\int c^2 u\,dx$ is the total energy of the system. Both are proportional to the "mass" $\int u\,dx$ which is a conserved quantity. In this case, we have

$$\mathcal{F}[u] = u_t + c u_x, \tag{134}$$

and the Fréchet derivative (81),

$$D\mathcal{F}[u]v = v_t + cv_x,$$ (135)

is not self-adjoint, as under the assumption of appropriate boundary conditions (e.g., periodic or homogeneous Dirichlet), integration by parts gives

$$\langle D\mathcal{F}[u]v, w \rangle = -\langle D\mathcal{F}[u]w, v \rangle.$$ (136)

It follows that there is no functional $\mathcal{A}[u]$ such that $D\mathcal{A} = \mathcal{F}$, i.e., $\mathcal{F}$ cannot be written as a variational problem.

4.2.1. Formal Lagrangian

The formal Lagrangian for the advection equation is obtained by multiplying (131) with the auxiliary variable $v(t, x)$. The solution vector of the extended system is denoted (u, v) with the corresponding section of the configuration bundle $\tilde{Y}$ denoted by $\tilde{\varphi}$. In coordinates, $\tilde{\varphi}$ is given explicitly as

$$\tilde{\varphi} : (t, x) \mapsto (t, x, u, v),$$ (137)

and its jet prolongation as

$$j^1\tilde{\varphi} : (t, x) \mapsto (t, x, u, v, u_t, u_x, v_t, v_x).$$ (138)

With that, the Lagrangian can be written as

$$L = \tilde{y}\left(z_t + cz_x\right),$$ (139)

with action functional

$$\mathcal{A}[\tilde{\varphi}] = \int L(j^1\tilde{\varphi}) dt\, dx = \int v\left(u_t + cu_x\right) dt\, dx.$$ (140)

The variation of the action with respect to the adjoint variable v gives the advection equation,

$$\frac{\delta\mathcal{A}[\tilde{\varphi}]}{\delta v} = u_t + cu_x = 0,$$ (141)

while the variation with respect to the original variable u yields the adjoint equation,

$$\frac{\delta\mathcal{A}[\tilde{\varphi}]}{\delta u} = -v_t - cv_x = 0,$$ (142)

always assuming vanishing variations at the boundaries. It is immediately observed that the adjoint equation has the same solution as the original equation, such that if u is a solution of the advection equation, then (u, u) solves the Euler–Lagrange equations of the extended Lagrangian (139). This means, that the advection equation is actually self-adjoint in the sense of Ibragimov, cf. Eq. (102), with $\lambda = -1$.

4.2.2. Continuous conservation laws

We will consider the conservation of mass and the L^2 norm of u. The general form of the Noether charge (101) is

$$\tilde{\mathcal{J}} = \int\left[\frac{\partial L}{\partial z_t^a}(j^1\tilde{\varphi}) \cdot \eta^a(\tilde{\varphi}) + \frac{\partial L}{\partial \tilde{z}_t^a}(j^1\tilde{\varphi}) \cdot \tilde{\eta}^a(\tilde{\varphi})\right] d^n\mathrm{x} = \int \tilde{J}^t(j^1\tilde{\varphi})\, d^n\mathrm{x} = \int v\eta\, d^n\mathrm{x},$$ (143)

with $\tilde{\mathcal{J}} = \text{const.}$ for all t. Since the advection equation is self-adjoint, we can use the identity as embedding map, i.e., $\phi(u) = u$, such that $v = u$, and the restriction of conservation laws for the extended system to the original system is straight forward.

The infinitesimal generator of the transformation that leads to mass conservation and its first jet prolongation are

$$V = \frac{\partial}{\partial y}, \qquad j^1V = \frac{\partial}{\partial y}.$$ (144)

Applying j^1V to (134), we obtain

$$j^1V(\mathcal{F}) = 0,$$ (145)

such that by (95) and (97) we find $\tilde{\eta} = 0$ and the extended vector field (94) is

$$\tilde{V} = \eta\frac{\partial}{\partial y} + \tilde{\eta}\frac{\partial}{\partial \tilde{y}} = \frac{\partial}{\partial y},$$ (146)

and its prolongation is

$$j^1\tilde{V} = \frac{\partial}{\partial y}.$$ (147)

The invariance of the Lagrangian (139) is easily confirmed, as

$$j^1 \tilde{V}(L) = \frac{\partial L}{\partial y} = 0. \tag{148}$$

The resulting conservation law (143) reads

$$\tilde{\mathcal{J}} = \int \tilde{J}^t (j^1 \tilde{\varphi}) \, dx = \int v \, dx = \text{const.} \quad \text{for all } t. \tag{149}$$

Upon restricting the Noether current $\tilde{J}$ with $v = \phi(u) = u$, this becomes conservation of the total mass in the system,

$$\mathcal{J} = \int (\tilde{J}^t \circ j^1 \Phi)(j^1 \varphi) \, dx = \int u \, dx = \text{const.} \quad \text{for all } t. \tag{150}$$

The conservation of momentum and energy follows exactly in the same way for $\eta = c$ and $\eta = \frac{1}{2}c^2$, respectively.

The infinitesimal generator corresponding to conservation of the L^2 norm is given by

$$V = y \frac{\partial}{\partial y}. \tag{151}$$

Applying its prolongation,

$$j^1 V = y \frac{\partial}{\partial y} + z_t \frac{\partial}{\partial z_t} + z_x \frac{\partial}{\partial z_x}, \tag{152}$$

to (131), we obtain

$$j^1 V(\mathcal{F}) = z_t + c z_x = \lambda \mathcal{F}, \tag{153}$$

with $\lambda = 1$, such that by (95) and (97) the extended vector field (94) becomes

$$\tilde{V} = y \frac{\partial}{\partial y} - \tilde{y} \frac{\partial}{\partial \tilde{y}}, \tag{154}$$

and its prolongation is

$$j^1 \tilde{V} = y \frac{\partial}{\partial y} + z_t \frac{\partial}{\partial y_t} + z_x \frac{\partial}{\partial y_x} - \tilde{y} \frac{\partial}{\partial \tilde{y}} - \tilde{z}_t \frac{\partial}{\partial \tilde{y}_t} - \tilde{z}_x \frac{\partial}{\partial \tilde{y}_x}. \tag{155}$$

The invariance condition of the Lagrangian (139),

$$j^1 \tilde{V}(L)(j^1 \tilde{\varphi}) = -v \left(u_t + c u_x \right) + v \left(u_t + c u_x \right) = 0, \tag{156}$$

is again fulfilled. The corresponding Noether charge (143) is

$$\tilde{\mathcal{J}} = \int \tilde{J}^t (j^1 \tilde{\varphi}) \, dx = \int vu \, dx. \tag{157}$$

Upon restricting the Noether current $\tilde{J}$ with $v = \phi(u) = u$, this becomes conservation of the L^2 norm,

$$\mathcal{J} = \int (\tilde{J}^t \circ j^1 \Phi)(j^1 \varphi) \, dx = \int u^2 \, dx = \text{const.} \quad \text{for all } t. \tag{158}$$

In the next sections we will derive the same conservation laws on the discrete level.

4.2.3. Midpoint discretisation

We discretise the Lagrangian (139) as described in Section 2.5 according to the midpoint rule (122), obtaining

$$L_d^{mp}(j^1 \tilde{\varphi}(\square)) = h_t h_x \frac{1}{4} \left(v_{\square^1} + v_{\square^2} + v_{\square^3} + v_{\square^4} \right) \left[\frac{1}{2} \left(\frac{u_{\square^4} - u_{\square^1}}{h_t} + \frac{u_{\square^3} - u_{\square^2}}{h_t} \right) + \frac{c}{2} \left(\frac{u_{\square^2} - u_{\square^1}}{h_x} + \frac{u_{\square^3} - u_{\square^4}}{h_x} \right) \right]. \tag{159}$$

In the shorthand notation defined in (125)–(127), this becomes

$$L_d(j^1 \tilde{\varphi}(\square)) = h_t h_x \, \overline{v}(\square) \left[\overline{u}_t(\square) + c \, \overline{u}_x(\square) \right], \tag{160}$$

which resembles the continuous Lagrangian.

Variational integrator

The discrete Euler–Lagrange field equations (59), corresponding to the variation of v_d, are computed as

$$0 = \frac{1}{4} \left[\frac{u_{i+1,j-1} - u_{i-1,j-1}}{2h_t} + 2 \frac{u_{i+1,j} - u_{i-1,j}}{2h_t} + \frac{u_{i+1,j+1} - u_{i-1,j+1}}{2h_t} \right]$$
$$+ \frac{c}{4} \left[\frac{u_{i-1,j+1} - u_{i-1,j-1}}{2h_x} + 2 \frac{u_{i,j+1} - u_{i,j-1}}{2h_x} + \frac{u_{i+1,j+1} - u_{i+1,j-1}}{2h_x} \right]. \tag{161}$$

As in the continuous case, the discrete adjoint equation has the exact same form as the discrete advection equation. We see that the derivatives are approximated by second-order centred finite differences. This means that we need initial data at two consecutive points in time, even though the advection equation is first order in time, so that initial data at one point in time should suffice. This problem is typical for applying the Veselov discretisation to formal Lagrangians. A simple solution will be provided in Section 4.3.4.

We also see that the time derivative is averaged in space and the spatial derivative is averaged in time. Under certain conditions, the spatial average of the time derivative can lead to grid-scale oscillations. This can be seen as follows. The time derivative in (161) features a spatial average of the form $\frac{1}{4}[1\ 2\ 1]$. So if on top of the actual solution there is some oscillation, e.g., of the form $[-1\ +1\ -1]$, this is eliminated by the average. Similarly, the spatial centred finite difference derivative, $\frac{1}{2h_x}[-1\ 0\ +1]$, "does not see" such an oscillation. For the case of periodic boundary conditions, the only way to prevent the instability to appear is to use an odd number of grid points in x. This is very similar to the phenomenon of checker-boarding often observed in simulations of the incompressible Euler equations.

In the next section we will prove some important discrete conservation laws of the resulting scheme. This exemplifies that even though a variational integrator has favourable conservation properties it might be unstable, and that not every discretisation of the Lagrangian leads to a good scheme.

Discrete conservation laws

For a formal Lagrangian of a single equation, the discrete symmetry condition (65) reads

$$
\begin{aligned}
0 = {}& \frac{\partial L_d}{\partial y_{\square^1}}\left(\tilde{\varphi}_{i,j}, \tilde{\varphi}_{i,j+1}, \tilde{\varphi}_{i+1,j+1}, \tilde{\varphi}_{i+1,j}\right) \cdot \eta_{i,j} + \frac{\partial L_d}{\partial \tilde{y}_{\square^1}}\left(\tilde{\varphi}_{i,j}, \tilde{\varphi}_{i,j+1}, \tilde{\varphi}_{i+1,j+1}, \tilde{\varphi}_{i+1,j}\right) \cdot \tilde{\eta}_{i,j} \\
& + \frac{\partial L_d}{\partial y_{\square^2}}\left(\tilde{\varphi}_{i,j}, \tilde{\varphi}_{i,j+1}, \tilde{\varphi}_{i+1,j+1}, \tilde{\varphi}_{i+1,j}\right) \cdot \eta_{i,j+1} + \frac{\partial L_d}{\partial \tilde{y}_{\square^2}}\left(\tilde{\varphi}_{i,j}, \tilde{\varphi}_{i,j+1}, \tilde{\varphi}_{i+1,j+1}, \tilde{\varphi}_{i+1,j}\right) \cdot \tilde{\eta}_{i,j+1} \\
& + \frac{\partial L_d}{\partial y_{\square^3}}\left(\tilde{\varphi}_{i,j}, \tilde{\varphi}_{i,j+1}, \tilde{\varphi}_{i+1,j+1}, \tilde{\varphi}_{i+1,j}\right) \cdot \eta_{i+1,j+1} + \frac{\partial L_d}{\partial \tilde{y}_{\square^3}}\left(\tilde{\varphi}_{i,j}, \tilde{\varphi}_{i,j+1}, \tilde{\varphi}_{i+1,j+1}, \tilde{\varphi}_{i+1,j}\right) \cdot \tilde{\eta}_{i+1,j+1} \\
& + \frac{\partial L_d}{\partial y_{\square^4}}\left(\tilde{\varphi}_{i,j}, \tilde{\varphi}_{i,j+1}, \tilde{\varphi}_{i+1,j+1}, \tilde{\varphi}_{i+1,j}\right) \cdot \eta_{i+1,j} + \frac{\partial L_d}{\partial \tilde{y}_{\square^4}}\left(\tilde{\varphi}_{i,j}, \tilde{\varphi}_{i,j+1}, \tilde{\varphi}_{i+1,j+1}, \tilde{\varphi}_{i+1,j}\right) \cdot \tilde{\eta}_{i+1,j}.
\end{aligned}
\tag{162}
$$

Here and in the following, evaluation of $\eta_{i,j}$ and $\tilde{\eta}_{i,j}$ at the fields $\varphi_{i,j}$ and $\tilde{\varphi}_{i,j}$ is implied. For the Lagrangian (159), this becomes

$$
\begin{aligned}
0 = {}& h_t h_x \frac{1}{8}\left[v_{i,j} + v_{i,j+1} + v_{i+1,j+1} + v_{i+1,j}\right]\left[\frac{\eta_{i+1,j} - \eta_{i,j}}{h_t} + \frac{\eta_{i+1,j+1} - \eta_{i,j+1}}{h_t}\right] \\
& + h_t h_x \frac{c}{8}\left[v_{i,j} + v_{i,j+1} + v_{i+1,j+1} + v_{i+1,j}\right]\left[\frac{\eta_{i,j+1} - \eta_{i,j}}{h_x} + \frac{\eta_{i+1,j+1} - \eta_{i+1,j}}{h_x}\right] \\
& + h_t h_x \frac{1}{8}\left[\tilde{\eta}_{i,j} + \tilde{\eta}_{i,j+1} + \tilde{\eta}_{i+1,j+1} + \tilde{\eta}_{i+1,j}\right]\left[\frac{u_{i+1,j} - u_{i,j}}{h_t} + \frac{u_{i+1,j+1} - u_{i,j+1}}{h_t}\right] \\
& + h_t h_x \frac{c}{8}\left[\tilde{\eta}_{i,j} + \tilde{\eta}_{i,j+1} + \tilde{\eta}_{i+1,j+1} + \tilde{\eta}_{i+1,j}\right]\left[\frac{u_{i,j+1} - u_{i,j}}{h_x} + \frac{u_{i+1,j+1} - u_{i+1,j}}{h_x}\right].
\end{aligned}
\tag{163}
$$

The corresponding discrete conservation law (72) is

$$
\begin{aligned}
\tilde{\mathcal{J}}_d = \sum_{j=1}^{N_1-1}\bigg[& \frac{\partial L_d}{\partial y_{\square^3}}\left(\tilde{\varphi}_{i,j}, \tilde{\varphi}_{i,j+1}, \tilde{\varphi}_{i+1,j+1}, \tilde{\varphi}_{i+1,j}\right) \cdot \eta_{i+1,j+1} + \frac{\partial L_d}{\partial y_{\square^4}}\left(\tilde{\varphi}_{i,j}, \tilde{\varphi}_{i,j+1}, \tilde{\varphi}_{i+1,j+1}, \tilde{\varphi}_{i+1,j}\right) \cdot \eta_{i+1,j} \\
& + \frac{\partial L_d}{\partial \tilde{y}_{\square^3}}\left(\tilde{\varphi}_{i,j}, \tilde{\varphi}_{i,j+1}, \tilde{\varphi}_{i+1,j+1}, \tilde{\varphi}_{i+1,j}\right) \cdot \tilde{\eta}_{i+1,j+1} + \frac{\partial L_d}{\partial \tilde{y}_{\square^4}}\left(\tilde{\varphi}_{i,j}, \tilde{\varphi}_{i,j+1}, \tilde{\varphi}_{i+1,j+1}, \tilde{\varphi}_{i+1,j}\right) \cdot \tilde{\eta}_{i+1,j}\bigg] = \text{const.}\quad\text{for all } i.
\end{aligned}
\tag{164}
$$

In contrast to the continuous case (143), here we also find contributions due to derivatives of the discrete Lagrangian L_d with respect to $\tilde{y}$. For the Lagrangian (159), the Noether charge becomes

$$
\begin{aligned}
\tilde{\mathcal{J}}_d = {}& h_t h_x \frac{1}{8}\sum_{j=1}^{N_1-1}\left[v_{i,j} + v_{i,j+1} + v_{i+1,j+1} + v_{i+1,j}\right]\left[\frac{\eta_{i+1,j+1} + \eta_{i+1,j}}{h_t} + c\,\frac{\eta_{i+1,j+1} - \eta_{i+1,j}}{h_x}\right] \\
& + h_t h_x \frac{1}{4}\sum_{j=1}^{N_1-1}\left[\tilde{\eta}_{i+1,j+1} + \tilde{\eta}_{i+1,j}\right]\left[\frac{1}{2}\left(\frac{u_{i+1,j} - u_{i,j}}{h_t} + \frac{u_{i+1,j+1} - u_{i,j+1}}{h_t}\right) + \frac{c}{2}\left(\frac{u_{i,j+1} - u_{i,j}}{h_x} + \frac{u_{i+1,j+1} - u_{i+1,j}}{h_x}\right)\right].
\end{aligned}
\tag{165}
$$

We begin with mass conservation. The discrete generator (c.f. (146)) is

$$
\eta_{i,j} = 1, \qquad \tilde{\eta}_{i,j} = 0.
\tag{166}
$$

The discrete symmetry condition (163) is fulfilled for (166). The second summation in (165) is zero, hence

$$
\tilde{\mathcal{J}}_d = h_x \sum_{j=1}^{N_1-1} \frac{1}{4}\left[v_{i,j} + v_{i,j+1} + v_{i+1,j+1} + v_{i+1,j}\right] = \text{const.}\quad\text{for all } i.
\tag{167}
$$

Identifying $v_{i,j} = u_{i,j}$, we obtain the discrete conservation law for the total mass in the system

$$\mathcal{J}_d = \frac{h_x}{4} \sum_{j=1}^{N_1-1} \left[u_{i,j} + u_{i,j+1} + u_{i+1,j+1} + u_{i+1,j} \right] = \text{const.} \quad \text{for all } i. \tag{168}$$

This expression supports grid oscillations. For the L^2 norm, the discrete generator (c.f. (154)) is

$$\eta_{i,j} = u_{i,j}, \qquad \tilde{\eta}_{i,j} = -v_{i,j}. \tag{169}$$

The discrete symmetry condition (163) is again easily checked to be fulfilled. The corresponding conservation law (72) is $\tilde{\mathcal{J}}_d = \text{const.}$ for all i with

$$\tilde{\mathcal{J}}_d = h_t h_x \frac{1}{8} \sum_{j=1}^{N_1-1} \left[v_{i,j} + v_{i,j+1} \right] \left[\frac{u_{i+1,j+1} + u_{i+1,j}}{h_t} + c \frac{u_{i+1,j+1} - u_{i+1,j}}{h_x} \right]$$

$$+ h_t h_x \frac{1}{8} \sum_{j=1}^{N_1-1} \left[v_{i+1,j+1} + v_{i+1,j} \right] \left[\frac{u_{i,j} + u_{i,j+1}}{h_t} - c \frac{u_{i,j+1} - u_{i,j}}{h_x} \right]. \tag{170}$$

Upon identifying $v_{i,j} = u_{i,j}$ and assuming periodic boundary conditions, the discrete Noether charge becomes

$$\mathcal{J}_d = h_x \sum_{j=1}^{N_1-1} u_{i,j} \left[\frac{u_{i+1,j-1} + 2u_{i+1,j} + u_{i+1,j+1}}{4} + \frac{ch_t}{2} \frac{u_{i+1,j+1} - u_{i+1,j-1}}{2h_x} \right]. \tag{171}$$

Just as the discrete mass (168), this expression also supports grid oscillations. It is somewhat unexpected that the discrete conservation law contains terms corresponding to the spatial component J^x of the Noether current. We therefore found an example where the discrete conserved quantity differs from the continuous conserved quantity. Nevertheless, $\mathcal{J}_d$ is consistent with the L^2 norm of u, as in the limit of $h_t \to 0$, holding c fixed, the additional term vanishes.

4.2.4. Trapezoidal discretisation

Writing out the discrete Lagrangian (123) for the linear advection equation, we obtain

$$L_d^{\text{tr}}\left(j^1 \tilde{\varphi}(\square)\right) = h_t h_x \left[\frac{1}{2} \left(\frac{v_{\square^1} + v_{\square^4}}{2} \frac{u_{\square^4} - u_{\square^1}}{h_t} + \frac{v_{\square^2} + v_{\square^3}}{2} \frac{u_{\square^3} - u_{\square^2}}{h_t} \right) + \frac{c}{2} \left(\frac{v_{\square^1} + v_{\square^2}}{2} \frac{u_{\square^2} - u_{\square^1}}{h_x} + \frac{v_{\square^3} + v_{\square^4}}{2} \frac{u_{\square^3} - u_{\square^4}}{h_x} \right) \right],$$

or in the compact notation of (129)–(130),

$$L_d^{\text{tr}}\left(j^1 \tilde{\varphi}(\square)\right) = h_t h_x \left[\overline{vu_t}(\square) + c \, \overline{vu_x}(\square) \right]. \tag{172}$$

Variational integrator

The discrete Euler–Lagrange equation (59) for the physical variable u_d reads

$$0 = \frac{u_{i+1,j} - u_{i-1,j}}{2h_t} + c \frac{u_{i,j+1} - u_{i,j-1}}{2h_x}, \tag{173}$$

which is the leapfrog scheme. The discrete Euler–Lagrange equation for the adjoint variable v_d takes exactly the same form.

Discrete conservation laws

The discrete symmetry condition (162) applied to (172) becomes

$$0 = h_t h_x \left[\frac{1}{2} \left(\frac{v_{i,j} + v_{i+1,j}}{2} \frac{\eta_{i+1,j} - \eta_{i,j}}{h_t} + \frac{v_{i,j+1} + v_{i+1,j+1}}{2} \frac{\eta_{i+1,j+1} - \eta_{i,j+1}}{h_t} \right) \right.$$

$$+ \frac{c}{2} \left(\frac{v_{i,j} + v_{i,j+1}}{2} \frac{\eta_{i,j+1} - \eta_{i,j}}{h_x} + \frac{v_{i+1,j+1} + v_{i+1,j}}{2} \frac{\eta_{i+1,j+1} - \eta_{i+1,j}}{h_x} \right)$$

$$+ \frac{1}{2} \left(\frac{\tilde{\eta}_{i,j} + \tilde{\eta}_{i+1,j}}{2} \frac{u_{i+1,j} - u_{i,j}}{h_t} + \frac{\tilde{\eta}_{i,j+1} + \tilde{\eta}_{i+1,j+1}}{2} \frac{u_{i+1,j+1} - u_{i,j+1}}{h_t} \right)$$

$$\left. + \frac{c}{2} \left(\frac{\tilde{\eta}_{i,j} + \tilde{\eta}_{i,j+1}}{2} \frac{u_{i,j+1} - u_{i,j}}{h_x} + \frac{\tilde{\eta}_{i+1,j+1} + \tilde{\eta}_{i+1,j}}{2} \frac{u_{i+1,j+1} - u_{i+1,j}}{h_x} \right) \right]. \tag{174}$$

The corresponding discrete Noether charge (164) is

$$\tilde{\mathcal{J}}_d = h_t h_x \frac{1}{2} \sum_{j=1}^{N_1-1} \left[\left(\frac{v_{i,j+1} + v_{i+1,j+1}}{2h_t} + c \frac{v_{i+1,j} + v_{i+1,j+1}}{2h_x} \right) \cdot \eta_{i+1,j+1} + \left(\frac{v_{i,j} + v_{i+1,j}}{2h_t} - c \frac{v_{i+1,j} + v_{i+1,j+1}}{2h_x} \right) \cdot \eta_{i+1,j} \right.$$

$$\left. + \left(\frac{u_{i+1,j+1} - u_{i,j+1}}{h_t} + c \frac{u_{i+1,j+1} - u_{i+1,j}}{h_x} \right) \cdot \tilde{\eta}_{i+1,j+1} + \left(\frac{u_{i+1,j} - u_{i,j}}{h_t} + c \frac{u_{i+1,j+1} - u_{i+1,j}}{h_x} \right) \cdot \tilde{\eta}_{i+1,j} \right]. \tag{175}$$

For the discrete generator of mass conservation (166), the symmetry condition (174) is fulfilled and the discrete Noether charge (175) amounts to (168). Similarly, the symmetry condition (174) is fulfilled for the discrete generator (169). The restriction of the corresponding discrete Noether charge (175) becomes

$$\mathcal{J}_d = \frac{h_x}{2} \sum_{j=1}^{N_1-1} \left[u_{i,j+1}\, u_{i+1,j+1} + u_{i,j}\, u_{i+1,j} \right] = \text{const.} \quad \text{for all } i, \tag{176}$$

which is indeed different from (171) and an immediate discretisation of the L^2 norm.

4.2.5. Simplified implicit scheme
The discrete Lagrangian (124) reads

$$L_d^{\mathrm{mt}}\big(\mathrm{j}^1\tilde\varphi(\square)\big) = h_t h_x\left[\frac{1}{2}\left(\frac{v_{\square^1} + v_{\square^4}}{2} \frac{u_{\square^4} - u_{\square^1}}{h_t} + \frac{v_{\square^2} + v_{\square^3}}{2} \frac{u_{\square^3} - u_{\square^2}}{h_t} \right) \right.$$
$$\left. + \frac{c}{8}\left(v_{\square^1} + v_{\square^2} + v_{\square^3} + v_{\square^4} \right)\left(\frac{u_{\square^2} - u_{\square^1}}{h_x} + \frac{u_{\square^3} - u_{\square^4}}{h_x} \right) \right],$$

or equivalently

$$L_d^{\mathrm{mt}}\big(\mathrm{j}^1\tilde\varphi(\square)\big) = h_t h_x\left[\overline{vu}_t(\square) + c\,\overline{v}(\square)\, u_x(\square) \right]. \tag{177}$$

Variational integrator

The discrete Euler–Lagrange equations (59) for the physical variable are obtained in the form

$$0 = \frac{u_{i+1,j} - u_{i-1,j}}{2h_t} + \frac{c}{4}\left[\frac{u_{i-1,j+1} - u_{i-1,j-1}}{2h_x} + 2\frac{u_{i,j+1} - u_{i,j-1}}{2h_x} + \frac{u_{i+1,j+1} - u_{i+1,j-1}}{2h_x} \right]. \tag{178}$$

As becomes already apparent by comparing the discrete Lagrangian (177) with (172) and (159), the time derivative is discretised in the same way as in the trapezoidal scheme (173) and the spatial derivative is discretised in the same way as in the midpoint scheme (161). Again, the discrete Euler–Lagrange equation for the adjoint variable takes the same form so that the system is self-adjoint.

Discrete conservation laws

In this case, the discrete symmetry condition (162) can be written as

$$0 = h_t h_x\left[\frac{1}{2}\left(\frac{v_{i,j} + v_{i+1,j}}{2} \frac{\eta_{i+1,j} - \eta_{i,j}}{h_t} + \frac{v_{i,j+1} + v_{i+1,j+1}}{2} \frac{\eta_{i+1,j+1} - \eta_{i,j+1}}{h_t} \right) \right.$$
$$+ \frac{c}{8}\left(v_{i,j} + v_{i,j+1} + v_{i+1,j+1} + v_{i+1,j} \right)\left(\frac{\eta_{i,j+1} - \eta_{i,j}}{h_x} + \frac{\eta_{i+1,j+1} - \eta_{i+1,j}}{h_x} \right)$$
$$+ \frac{1}{2}\left(\frac{\tilde\eta_{i,j} + \tilde\eta_{i+1,j}}{2} \frac{u_{i+1,j} - u_{i,j}}{h_t} + \frac{\tilde\eta_{i,j+1} + \tilde\eta_{i+1,j+1}}{2} \frac{u_{i+1,j+1} - u_{i,j+1}}{h_t} \right)$$
$$\left. + \frac{c}{8}\left(\tilde\eta_{i,j} + \tilde\eta_{i,j+1} + \tilde\eta_{i+1,j+1} + \tilde\eta_{i+1,j} \right)\left(\frac{u_{i,j+1} - u_{i,j}}{h_x} + \frac{u_{i+1,j+1} - u_{i+1,j}}{h_x} \right) \right], \tag{179}$$

which yields the discrete Noether charge, (c.f., (164)),

$$\tilde{\mathcal{J}}_d = h_t h_x \frac{1}{2} \sum_{j=1}^{N_1-1}\left[\left(\frac{v_{i,j+1} + v_{i+1,j+1}}{2h_t} + c\, \frac{v_{i,j} + v_{i,j+1} + v_{i+1,j} + v_{i+1,j+1}}{4h_x} \right)\cdot \eta_{i+1,j+1} \right.$$
$$+ \left(\frac{v_{i,j} + v_{i+1,j}}{2h_t} - c\, \frac{v_{i,j} + v_{i,j+1} + v_{i+1,j} + v_{i+1,j+1}}{4h_x} \right)\cdot \eta_{i+1,j}$$
$$+ \left(\frac{u_{i+1,j+1} - u_{i,j+1}}{h_t} + \frac{c}{4}\frac{u_{i,j+1} - u_{i,j}}{h_x} + \frac{c}{4}\frac{u_{i+1,j+1} - u_{i+1,j}}{h_x} \right)\cdot \tilde\eta_{i+1,j+1}$$
$$\left. + \left(\frac{u_{i+1,j} - u_{i,j}}{h_t} + \frac{c}{4}\frac{u_{i,j+1} - u_{i,j}}{h_x} + \frac{c}{4}\frac{u_{i+1,j+1} - u_{i+1,j}}{h_x} \right)\cdot \tilde\eta_{i+1,j} \right]. \tag{180}$$

The symmetry condition (179) is satisfied for both, mass conservation and the L^2 norm. The corresponding discrete conservation laws reduce to (168) and (176), respectively.

4.2.6. Dissipation and dispersion
In this section we want to carry out a simple von Neumann analysis of some of the variational integrators we have derived in order to analyse their dissipation and dispersion behaviour.

The oscillatory function $u(t, x) = e^{-\iota\omega t + \iota k x}$ is a solution of the advection equation (131) if and only if (ω, k) satisfies the exact dispersion relation $\omega = ck$. When collocated at grid points $t_i = ih_t$, $x_j = jh_x$, this amounts to

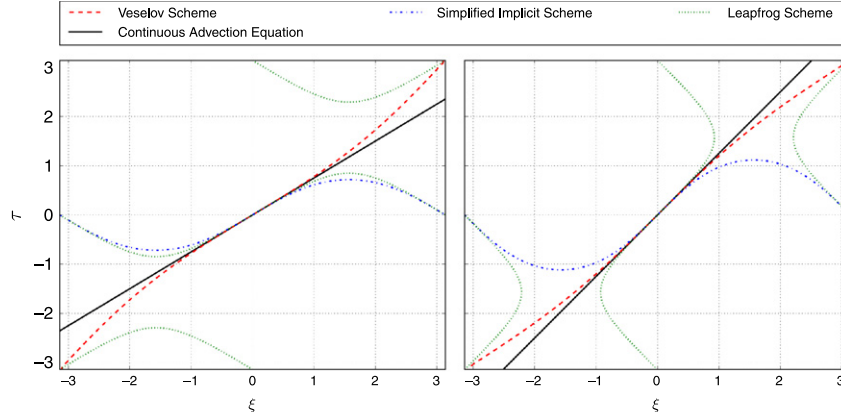$$u_{ij} = e^{-\iota(\tau i - \xi j)}, \tag{181}$$

Fig. 5. Numerical dispersion relation of the Veselov scheme (red dashed line), the simplified implicit scheme (blue dash-dotted line) and the leap frog scheme (green dotted line) applied to the linear advection equation with $c/c_{\text{grid}} = 0.75$ (left) and $c/c_{\text{grid}} = 1.25$ (right). For comparison the exact dispersion relation is also shown (black solid line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

with $\tau = \omega h_t$ and $\xi = k h_x$. In analogy with the continuous case, the collocated plane wave u_{ij} is a solution of the numerical scheme if and only if (τ, ξ) satisfies the numerical dispersion relation. For the case of the Veselov scheme (161), the dispersion relation can be written as

$$\sin(\tau)\big[1 + \cos(\xi)\big] - \frac{c}{c_{\text{grid}}}\big[1 + \cos(\tau)\big]\sin(\xi) = 0, \tag{182}$$

where $c_{\text{grid}} = h_x/h_t$. This is the same dispersion relation as the one obtained by Ascher and McLachlan [47] for the box scheme. For comparison, we consider the simplified implicit scheme (178), for which the dispersion relation is

$$\sin(\tau) - \frac{1}{2}\frac{c}{c_{\text{grid}}}\big[1 + \cos(\tau)\big]\sin(\xi) = 0, \tag{183}$$

as well as the leap frog scheme (173), for which the dispersion relation is

$$\sin(\tau) - \frac{c}{c_{\text{grid}}}\sin(\xi) = 0. \tag{184}$$

One can observe that in all cases the left-hand side is real valued, i.e., numerical dissipation is exactly zero, but there is numerical dispersion. Fig. 5 shows the solution of the numerical dispersion relation in comparison with the exact linear dispersion of the advection equation in the (ξ, τ)-plane for different values of the ratio c/c_{grid}. As usual, for (ξ, τ) near the origin, the numerical dispersion is a good approximation of the exact linear relation, and thus the corresponding harmonics have a phase velocity close to the exact value c. On the other hand, when the wave number is larger than half of the grid wave number $1/h_x$, the frequency of the harmonic deviates from the exact dispersion relation. The corresponding harmonics have a phase velocity different from the exact value c, leading to numerical dispersion.

For the simplified implicit scheme, one can also observe local maxima of the numerical dispersion, which correspond to a zero numerical group velocity, while the exact value for the advection equation is equal to the phase velocity c.

The left-hand-side panel of Fig. 5 shows the dispersion relation for the case $c/c_{\text{grid}} < 1$. We find that the leapfrog scheme supports parasitic modes, that is to each $\xi \in [-\pi, +\pi]$ there are two values of τ. The high-frequency branches are called parasitic modes. The Veselov and the simplified implicit scheme also feature such modes at the $\tau = -\pi$ and $\tau = +\pi$ lines. As we will see in the numerical experiments, these branches will be populated only if the spectrum of the solution is sufficiently broad to cover the nonlinear region of the dispersion relation, that is for waves which are not well represented by the chosen grid parameters.

For didactical purposes, we have also considered the case $c/c_{\text{grid}} > 1$ on the right-hand-side panel of Fig. 5. One can see that the numerical group velocity of the leapfrog scheme becomes infinite, e.g., for $\xi = \pm\sin^{-1}(0.8)$, which renders the scheme unstable. This is always observed for $c/c_{\text{grid}} > 1$ which corresponds to the Courant–Friedrichs–Levy (CFL) stability criterion that limits the timestep for explicit schemes. The implicit variational integrators do not have such a limitation. In the following examples, the timestep is chosen small enough so that these instabilities do not appear.

4.2.7. Numerical examples

We consider two test cases, a sum of cosines and a Gaussian. The first is purely meant to verify the dispersion relation, whereas the second is used to judge the quality of the solution. In both cases, the domain is $\Omega = [-0.5, +0.5)$ with periodic boundary conditions, the timestep is $h_t = 2.5 \times 10^{-3}$, the phase velocity is set to $c = 1$, and the number of points in space and time is $n_x = 255$ and $n_t = 4000$, respectively. Hence, the time interval is $[0, 10]$, which corresponds to ten passes of the wave packet through the domain.

The following tests are implemented in Python [48,49] using NumPy [50] and the sparse linear algebra module of SciPy [51]. Visualisation was done using matplotlib [52,53].

Experimental dispersion relation

In this example, we initialise u with a sum of cosines,

$$u_{C,0}(x) = \sum_{i=1}^{n_x/2} \cos(i2\pi x), \tag{185}$$

(a) Veselov scheme. (b) Simplified implicit scheme. (c) Leapfrog scheme.
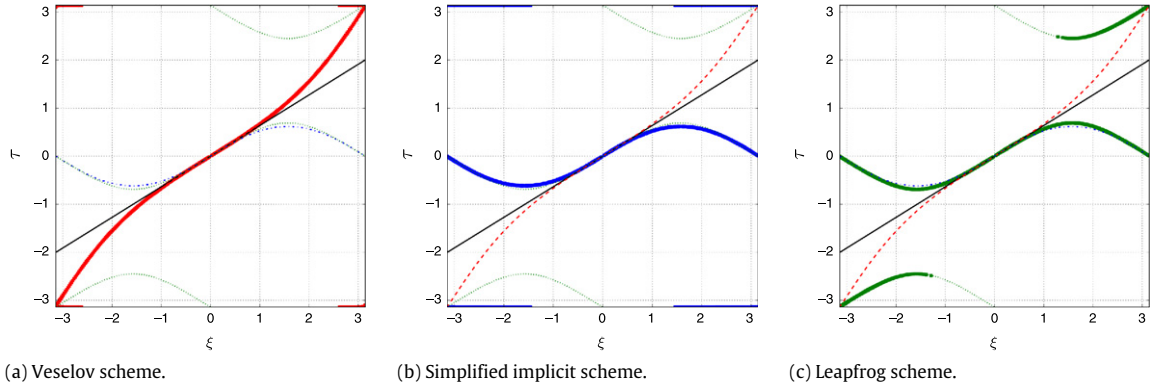
Fig. 6. Theoretical and experimental dispersion relation for the sum of cosines. The experimental values lie right on top of the theoretical curve. Red dashed line: Veselov scheme. Blue dash-dotted line: Simplified implicit scheme. Green dotted line: Leapfrog scheme. Black solid line: Analytical dispersion relation.

which excites all the modes supported by the grid in order to experimentally verify the dispersion relations (182)–(184) from the previous section. Fig. 6 shows the theoretical and experimental dispersion relations for the Veselov scheme, the simplified implicit scheme and the leapfrog scheme. The theoretical dispersion relations are well matched by the simulation, including some of the parasitic modes.

Gaussian wave

In this section, we consider a Gaussian,

$$u_{G,0}(x) = \frac{1}{\sigma\sqrt{2\pi}} \, \exp\left(-\frac{1}{2}\left[\frac{x}{\sigma}\right]^2\right), \tag{186}$$

with $\sigma = 0.1$, so that the spectrum is well within the linear region of the dispersion relation. Fig. 7 shows that after 10 passes the form of the Gaussian is close to the initial one for both, the simplified implicit variational integrator (178) and the leapfrog scheme (173), and seems to exactly match the initial conditions for the Veselov scheme (161). The leapfrog scheme shows slightly less distortion than the simplified implicit scheme, which is explained by their dispersion relations, which is closer to the analytic dispersion relation for the leapfrog scheme and even more so for the Veselov scheme. For all three integrators, conservation of energy and the L^2 norm is also shown in Fig. 7.

4.3. Vorticity equation

Next we consider an example with more than one equation and more than one spatial dimension, the vorticity equation,

$$\omega_t + \{\psi, \omega\} = 0, \qquad \omega = \Delta\psi. \tag{187}$$

Here, ψ is the streaming function and ω is the vorticity. The vorticity equation arises by computing the curl of the incompressible Euler equation in two dimensions,

$$u_t + (u \cdot \nabla)u + \nabla p = 0, \tag{188}$$

where u is the fluid velocity and p the pressure. Specifically, ψ and ω are related to the velocity u by $u = \nabla_\perp \psi$ and $\omega = \nabla_\perp \cdot u$ with $\nabla_\perp = (-\partial_y, \partial_x)^T$. The Poisson bracket $\{\cdot, \cdot\}$ is defined by

$$\{\psi, \omega\} = \psi_x \omega_y - \psi_y \omega_x, \tag{189}$$

which coincides with the determinant of the Jacobian of the transformation $(x, y) \mapsto (\psi, \omega)$. For this reason, the 2-dimensional Poisson bracket is also referred to as Jacobian. The variational discretisation of (187) leads to a very interesting result, namely Arakawa's famous discretisation of the Poisson bracket [54].

4.3.1. Formal Lagrangian

The formal Lagrangian of the vorticity equation (187) is

$$L(j^2\tilde{\varphi}) = \zeta\left(\omega_t + \{\psi, \omega\}\right) + \chi\left(\omega - \Delta\psi\right), \tag{190}$$

with solution vector $(\omega, \psi, \zeta, \chi)$. Here, the formal Lagrangian L is defined on the second jet-bundle $J^2\tilde{Y}$ due to the Laplace operator. In order to avoid second order derivatives, we apply a symmetrisation (corresponding to an integration by parts in the action functional) in the term that produces the Laplacian. This yields an equivalent Lagrangian, which however is defined on the first jet-bundle $J^1\tilde{Y}$.

Some care has to be taken when discretising the Poisson bracket (see Salmon and Talley [55]). In order to retain the antisymmetry property of the continuous bracket at the discrete level, a symmetrisation has to be introduced in the Lagrangian. Using integration by
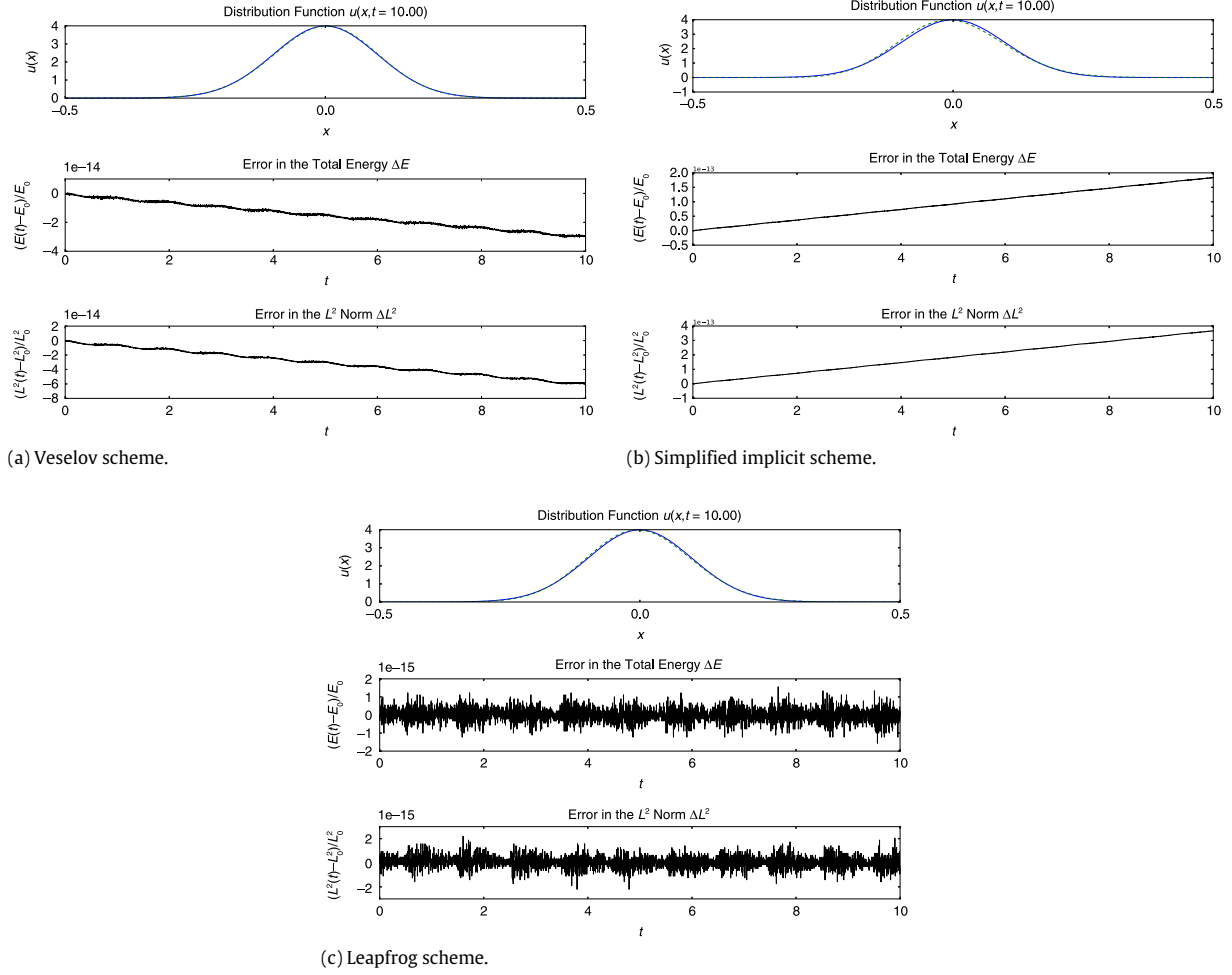
(a) Veselov scheme.

(b) Simplified implicit scheme.

(c) Leapfrog scheme.

Fig. 7. Gaussian after ten passes with the Veselov scheme, the simplified implicit scheme and the leapfrog scheme. Comparing the solution (green dashed curve) to the initial condition (blue curve) the effect of dispersion becomes visible. Energy and the L^2 norm are well preserved. For the Veselov scheme as well as the simplified implicit scheme there is some drift in the errors, but this is extremely small.

parts while assuming appropriate boundary conditions, it is seen that the cyclic permutations of the functions in the integrand are all identical,[4]

$$\int \zeta \{\psi, \omega\} \, dx \, dy = \int \psi \{\omega, \zeta\} \, dx \, dy = \int \omega \{\zeta, \psi\} \, dx \, dy. \tag{191}$$

Instead of selecting one of those equivalent forms, a convex combination can be considered, namely,

$$\int \zeta \{\psi, \omega\} \, dx \, dy = \int \left[\alpha \, \zeta \{\psi, \omega\} + \beta \, \psi \{\omega, \zeta\} + \gamma \, \omega \{\zeta, \psi\} \right] dx \, dy, \tag{192}$$

with $\alpha + \beta + \gamma = 1$. One finds that the symmetric case, $\alpha = \beta = \gamma = 1/3$, is the one that retains the properties of the bracket at the discrete level. We therefore use the modified Lagrangian

$$L'(j^1 \tilde{\varphi}) = \zeta \omega_t + \frac{1}{3} \left[\zeta \{\psi, \omega\} + \psi \{\omega, \zeta\} + \omega \{\zeta, \psi\} \right] + \chi \omega + \nabla \chi \cdot \nabla \psi. \tag{193}$$

Computing the variational derivatives of the action with respect to the adjoint variables, we obtain (187). The variational derivatives with respect to the physical variables give

$$\zeta_t + \{\psi, \zeta\} = \chi, \qquad \Delta \chi = \{\omega, \zeta\}. \tag{194}$$

This extended system of equations is obviously not self-adjoint, but we can still find a simple compatible solution of the adjoint equations, namely $(\zeta, \chi) = \phi(\omega, \psi) = (\omega, 0)$. Assume we can set $\zeta = \omega$, then the Poisson bracket in the adjoint Poisson equation vanishes, which

---

[4] This holds assuming boundary conditions such that the surface terms of the integration vanish, e.g., periodic or homogeneous Dirichlet.

implies that χ is a harmonic function. In particular, we can choose $\chi = 0$, so that the right-hand side of the equation for ζ becomes zero and ζ fulfils the same equation as ω. This justifies the choice of ϕ.

4.3.2. Continuous conservation laws

In order to determine the conservation laws for the vorticity equation, it is not enough to just consider vertical transformations as in (2.3). We have to consider also horizontal transformations, as the interesting symmetries of the vorticity equation are not generated by purely vertical vector fields. To do so, the discrete Noether theorem as it is presented in Section 2.6 is not sufficiently general. Unfortunately, it is not straight forward to consider symmetry generators with horizontal components in the framework of finite difference discretisations of the Lagrangian as we use it here. For that reason, we follow a simplified approach to determine the continuous conservation laws for the vorticity equation which is amenable to discretisation. A rigorous analysis of the continuous and discrete conservation laws for the vorticity equation is postponed to a subsequent exposition.

In the following, we will consider the vorticity equation for ω alone,

$$\mathcal{F}[\omega] = \omega_t + \{\psi, \omega\} = 0, \tag{195}$$

without the Poisson equation, assuming a constant-in-time streaming function ψ. The Lagrangian for this system reads

$$\bar{L} = \tilde{y}(z_t + \psi_x z_y - \psi_y z_x). \tag{196}$$

It is easily seen that Euler–Lagrange equations (10) for (196) amount to (195) and

$$\zeta_t + \{\psi, \zeta\} = 0, \tag{197}$$

such that the extended system is self-adjoint and we can construct a solution for ζ simply by $\phi(\omega) = \omega$. The generating vector field for vertical transformations of solutions of the vorticity equation (195) and its first jet prolongation are written as

$$V = \eta \frac{\partial}{\partial y} \quad \text{and} \quad j^1 V = \eta \frac{\partial}{\partial y} + \eta_\mu \frac{\partial}{\partial z_\mu}. \tag{198}$$

The action of a general vector field V on $\mathcal{F}$ is given by

$$j^1 V(\mathcal{F}) = \eta_t + \psi_x \eta_y - \psi_y \eta_x. \tag{199}$$

This is used to determine the component $\tilde{\eta}$ of the corresponding extended vector field $\tilde{V}$ and its first jet prolongation $j^1\tilde{V}$,

$$\tilde{V} = \eta \frac{\partial}{\partial y} + \tilde{\eta} \frac{\partial}{\partial \tilde{y}} \quad \text{and} \quad j^1 \tilde{V} = \eta \frac{\partial}{\partial y} + \eta_\mu \frac{\partial}{\partial z_\mu} + \tilde{\eta} \frac{\partial}{\partial \tilde{y}} + \tilde{\eta}_\mu \frac{\partial}{\partial \tilde{z}_\mu}. \tag{200}$$

Its action on the Lagrangian (196) is

$$j^1 \tilde{V}(\bar{L}) = \tilde{y} (\eta_t + \psi_x \eta_y - \psi_y \eta_x) + \tilde{\eta} (z_t + \psi_x z_y - \psi_y z_x). \tag{201}$$

For conservation of circulation ("mass") we have $\eta = 1$, and therefore

$$j^1 V = \frac{\partial}{\partial y}. \tag{202}$$

The action of the prolonged vector field on (195) becomes

$$j^1 V(\mathcal{F}) = 0, \tag{203}$$

so that $\tilde{\eta} = 0$ and

$$j^1 \tilde{V} = \frac{\partial}{\partial y}. \tag{204}$$

The action of the prolonged vector field on the Lagrangian (196) is

$$j^1 \tilde{V}(\bar{L}) = \frac{\partial \bar{L}}{\partial y} = 0, \tag{205}$$

that is we have a symmetry, and the corresponding conservation law (Noether charge) reads

$$\tilde{\mathcal{J}} = \int \tilde{J}^t \, dx = \int \frac{\partial L}{\partial z_t} (j^1 \tilde{\varphi}) \cdot \eta(\tilde{\varphi}) \, dx = \int \zeta \, dx = \text{const. for all } t. \tag{206}$$

Upon restricting the Noether current $\tilde{J}$ with $\zeta = \phi(\omega) = \omega$, this becomes conservation of the total circulation in the system,

$$\mathcal{J} = \int (\tilde{J}^t \circ j^1\Phi)(j^1\varphi)\, dx = \int \omega\, dx = \text{const.} \quad \text{for all } t. \tag{207}$$

As here, ψ is treated as constant in time, energy conservation is obtained in the same way for $\eta = \frac{1}{2}\psi$.

Next we consider enstrophy, for which $\eta = y$, so that the first jet prolongation of the generating vector field,

$$j^1 V = y\frac{\partial}{\partial y} + z_t\frac{\partial}{\partial z_t} + z_x\frac{\partial}{\partial z_x} + z_y\frac{\partial}{\partial z_y}, \tag{208}$$

acts on the vorticity equation (196) as

$$j^1 V(\mathcal{F}) = z_t + \psi_x z_y - \psi_y z_x = \mathcal{F}, \tag{209}$$

hence $\tilde{\eta} = -\tilde{y}$ and

$$j^1 \tilde{V} = y\frac{\partial}{\partial y} + z_t\frac{\partial}{\partial z_t} + z_x\frac{\partial}{\partial z_x} + z_y\frac{\partial}{\partial z_y} - \tilde{y}\frac{\partial}{\partial \tilde{y}} - \tilde{z}_t\frac{\partial}{\partial \tilde{z}_t} - \tilde{z}_x\frac{\partial}{\partial \tilde{z}_x} - \tilde{z}_y\frac{\partial}{\partial \tilde{z}_y}. \tag{210}$$

The prolongation of the extended vector field acts on the Lagrangian (196) as

$$j^1 \tilde{V}(\bar{L}) = \tilde{y}\,(z_t + \psi_x z_y - \psi_y z_x) - \tilde{y}\,(z_t + \psi_x z_y - \psi_y z_x) = 0, \tag{211}$$

so that we have a symmetry whose Noether charge is computed as

$$\tilde{\mathcal{J}} = \int \tilde{J}^t\, dx = \int \frac{\partial L}{\partial z_t}(j^1\tilde{\varphi}) \cdot \eta(\tilde{\varphi})\, dx = \int \zeta\omega\, dx = \text{const. for all } t. \tag{212}$$

Restricting the Noether current $\tilde{J}$ with $\zeta = \phi(\omega) = \omega$, this becomes conservation of enstrophy,

$$\mathcal{J} = \int (\tilde{J}^t \circ j^1\Phi)(j^1\varphi)\, dx = \int \omega^2\, dx = \text{const.} \quad \text{for all } t. \tag{213}$$

In summary, we obtain conservation of

(a) circulation ("mass")

$$\int \omega\, dx\, dy, \tag{214a}$$

(b) enstrophy (L^2 norm)

$$\int \omega^2\, dx\, dy, \tag{214b}$$

(c) and kinetic energy

$$\frac{1}{2}\int \psi\omega\, dx\, dy. \tag{214c}$$

For the symmetrised Lagrangian, c.f. Eq. (193),

$$\bar{L}'(j^1\tilde{\varphi}) = \zeta\omega_t + \frac{1}{3}\Big[\zeta\{\psi, \omega\} + \psi\{\omega, \zeta\} + \omega\{\zeta, \psi\}\Big], \tag{215}$$

we obtain the same conservation laws, although in that case they correspond to divergence symmetries.

4.3.3. Variational integrator

To discretise the Lagrangian $L(\tilde{\varphi}, \tilde{\varphi}_t, \tilde{\varphi}_x, \tilde{\varphi}_y)$ from (193) we adopt the same strategy as in Section 4.2.5. We use simple finite differences to approximate the derivatives (see Fig. 8), the midpoint rule for the time integral and the trapezoidal rule for the two spatial integrals,

$$
\begin{aligned}
L_d\big(j^1\tilde{\varphi}(\square)\big) =\ & \frac{h_t h_x h_y}{4} L'\Big(\tfrac{\tilde{\varphi}_{\square 1}+\tilde{\varphi}_{\square 5}}{2},\ \tfrac{\tilde{\varphi}_{\square 5}-\tilde{\varphi}_{\square 1}}{h_t},\ \tfrac{1}{2}\Big(\tfrac{\tilde{\varphi}_{\square 2}-\tilde{\varphi}_{\square 1}}{h_x}+\tfrac{\tilde{\varphi}_{\square 6}-\tilde{\varphi}_{\square 5}}{h_x}\Big),\ \tfrac{1}{2}\Big(\tfrac{\tilde{\varphi}_{\square 4}-\tilde{\varphi}_{\square 1}}{h_y}+\tfrac{\tilde{\varphi}_{\square 8}-\tilde{\varphi}_{\square 5}}{h_y}\Big)\Big) \\
& + \frac{h_t h_x h_y}{4} L'\Big(\tfrac{\tilde{\varphi}_{\square 2}+\tilde{\varphi}_{\square 6}}{2},\ \tfrac{\tilde{\varphi}_{\square 6}-\tilde{\varphi}_{\square 2}}{h_t},\ \tfrac{1}{2}\Big(\tfrac{\tilde{\varphi}_{\square 2}-\tilde{\varphi}_{\square 1}}{h_x}+\tfrac{\tilde{\varphi}_{\square 6}-\tilde{\varphi}_{\square 5}}{h_x}\Big),\ \tfrac{1}{2}\Big(\tfrac{\tilde{\varphi}_{\square 3}-\tilde{\varphi}_{\square 2}}{h_y}+\tfrac{\tilde{\varphi}_{\square 7}-\tilde{\varphi}_{\square 6}}{h_y}\Big)\Big) \\
& + \frac{h_t h_x h_y}{4} L'\Big(\tfrac{\tilde{\varphi}_{\square 3}+\tilde{\varphi}_{\square 7}}{2},\ \tfrac{\tilde{\varphi}_{\square 7}-\tilde{\varphi}_{\square 3}}{h_t},\ \tfrac{1}{2}\Big(\tfrac{\tilde{\varphi}_{\square 3}-\tilde{\varphi}_{\square 4}}{h_x}+\tfrac{\tilde{\varphi}_{\square 7}-\tilde{\varphi}_{\square 8}}{h_x}\Big),\ \tfrac{1}{2}\Big(\tfrac{\tilde{\varphi}_{\square 3}-\tilde{\varphi}_{\square 2}}{h_y}+\tfrac{\tilde{\varphi}_{\square 7}-\tilde{\varphi}_{\square 6}}{h_y}\Big)\Big) \\
& + \frac{h_t h_x h_y}{4} L'\Big(\tfrac{\tilde{\varphi}_{\square 4}+\tilde{\varphi}_{\square 8}}{2},\ \tfrac{\tilde{\varphi}_{\square 8}-\tilde{\varphi}_{\square 4}}{h_t},\ \tfrac{1}{2}\Big(\tfrac{\tilde{\varphi}_{\square 3}-\tilde{\varphi}_{\square 4}}{h_x}+\tfrac{\tilde{\varphi}_{\square 7}-\tilde{\varphi}_{\square 8}}{h_x}\Big),\ \tfrac{1}{2}\Big(\tfrac{\tilde{\varphi}_{\square 4}-\tilde{\varphi}_{\square 1}}{h_y}+\tfrac{\tilde{\varphi}_{\square 8}-\tilde{\varphi}_{\square 5}}{h_y}\Big)\Big).
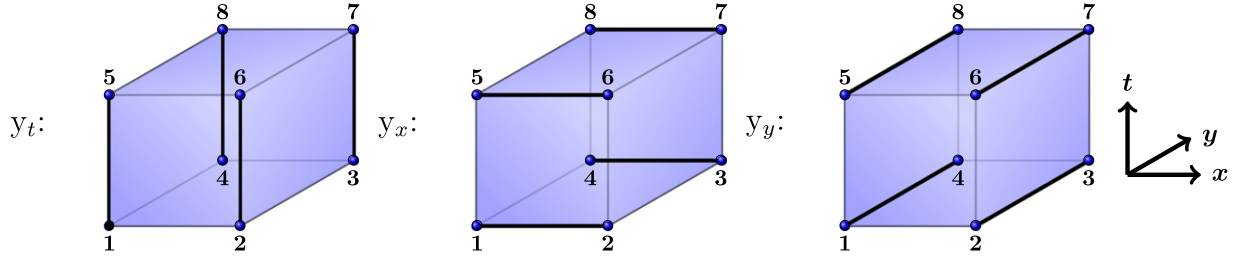\end{aligned}
\tag{216}
$$

Fig. 8. For a given spacetime grid cell, there are four possible ways of defining derivatives in the different coordinate directions (t, x, y), highlighted along the black lines.

In $2+1$ dimensions, the discrete Euler–Lagrange field equations have eight contributions instead of four,

$$
0 = \frac{\partial L_d}{\partial \tilde{\varphi}_{\square 1}^a} \Big(\tilde{\varphi}_{i,j,k}, \tilde{\varphi}_{i,j+1,k}, \tilde{\varphi}_{i,j+1,k+1}, \tilde{\varphi}_{i,j,k+1}, \tilde{\varphi}_{i+1,j,k}, \tilde{\varphi}_{i+1,j+1,k}, \tilde{\varphi}_{i+1,j+1,k+1}, \tilde{\varphi}_{i+1,j,k+1} \Big)
$$

$$
+ \frac{\partial L_d}{\partial \tilde{\varphi}_{\square 2}^a} \Big(\tilde{\varphi}_{i,j-1,k}, \tilde{\varphi}_{i,j,k}, \tilde{\varphi}_{i,j,k+1}, \tilde{\varphi}_{i,j-1,k+1}, \tilde{\varphi}_{i+1,j-1,k}, \tilde{\varphi}_{i+1,j,k}, \tilde{\varphi}_{i+1,j,k+1}, \tilde{\varphi}_{i+1,j-1,k+1} \Big)
$$

$$
+ \frac{\partial L_d}{\partial \tilde{\varphi}_{\square 3}^a} \Big(\tilde{\varphi}_{i,j-1,k-1}, \tilde{\varphi}_{i,j,k-1}, \tilde{\varphi}_{i,j,k}, \tilde{\varphi}_{i,j-1,k}, \tilde{\varphi}_{i+1,j-1,k-1}, \tilde{\varphi}_{i+1,j,k-1}, \tilde{\varphi}_{i+1,j,k}, \tilde{\varphi}_{i+1,j-1,k} \Big)
$$

$$
+ \frac{\partial L_d}{\partial \tilde{\varphi}_{\square 4}^a} \Big(\tilde{\varphi}_{i,j,k-1}, \tilde{\varphi}_{i,j+1,k-1}, \tilde{\varphi}_{i,j+1,k}, \tilde{\varphi}_{i,j,k}, \tilde{\varphi}_{i+1,j,k-1}, \tilde{\varphi}_{i+1,j+1,k-1}, \tilde{\varphi}_{i+1,j+1,k}, \tilde{\varphi}_{i+1,j,k} \Big)
$$

$$
+ \frac{\partial L_d}{\partial \tilde{\varphi}_{\square 5}^a} \Big(\tilde{\varphi}_{i-1,j,k}, \tilde{\varphi}_{i-1,j+1,k}, \tilde{\varphi}_{i-1,j+1,k+1}, \tilde{\varphi}_{i-1,j,k+1}, \tilde{\varphi}_{i,j,k}, \tilde{\varphi}_{i,j+1,k}, \tilde{\varphi}_{i,j+1,k+1}, \tilde{\varphi}_{i,j,k+1} \Big)
$$

$$
+ \frac{\partial L_d}{\partial \tilde{\varphi}_{\square 6}^a} \Big(\tilde{\varphi}_{i-1,j-1,k}, \tilde{\varphi}_{i-1,j,k}, \tilde{\varphi}_{i-1,j,k+1}, \tilde{\varphi}_{i-1,j-1,k+1}, \tilde{\varphi}_{i,j-1,k}, \tilde{\varphi}_{i,j,k}, \tilde{\varphi}_{i,j,k+1}, \tilde{\varphi}_{i,j-1,k+1} \Big)
$$

$$
+ \frac{\partial L_d}{\partial \tilde{\varphi}_{\square 7}^a} \Big(\tilde{\varphi}_{i-1,j-1,k-1}, \tilde{\varphi}_{i-1,j,k-1}, \tilde{\varphi}_{i-1,j,k}, \tilde{\varphi}_{i-1,j-1,k}, \tilde{\varphi}_{i,j-1,k-1}, \tilde{\varphi}_{i,j,k-1}, \tilde{\varphi}_{i,j,k}, \tilde{\varphi}_{i,j-1,k} \Big)
$$

$$
+ \frac{\partial L_d}{\partial \tilde{\varphi}_{\square 8}^a} \Big(\tilde{\varphi}_{i-1,j,k-1}, \tilde{\varphi}_{i-1,j+1,k-1}, \tilde{\varphi}_{i-1,j+1,k}, \tilde{\varphi}_{i-1,j,k}, \tilde{\varphi}_{i,j,k-1}, \tilde{\varphi}_{i,j+1,k-1}, \tilde{\varphi}_{i,j+1,k}, \tilde{\varphi}_{i,j,k} \Big). \tag{217}
$$

The resulting integrator for the vorticity equation takes the form

$$
\frac{\omega_{i+1} - \omega_{i-1}}{2h_t} + \frac{1}{8} \Big[A(\psi_{i+1}, \omega_{i+1}) + A(\psi_i, \omega_{i+1}) + A(\psi_{i+1}, \omega_i) + 2A(\psi_i, \omega_i) + A(\psi_{i-1}, \omega_i) + A(\psi_i, \omega_{i-1}) + A(\psi_{i-1}, \omega_{i-1}) \Big], \tag{218}
$$

where $\omega_i = (\omega_{ij})_j$ and $\psi_i = (\psi_{ij})_j$ are the rows of the solution matrix corresponding to constant-time slices, whereas A denotes Arakawa's discretisation of the Poisson bracket [54], given by

$$
A(\psi, \omega) = \frac{1}{3} \Big(A^{++}(\psi, \omega) + A^{+\times}(\psi, \omega) + A^{\times+}(\psi, \omega) \Big), \tag{219}
$$

with ω and ψ being row vectors and

$$
A^{++}(\psi, \omega) = \frac{1}{4h_x h_y} \Big[\big(\psi_{+0} - \psi_{-0} \big) \big(\omega_{0+} - \omega_{0-} \big) - \big(\psi_{0+} - \psi_{0-} \big) \big(\omega_{+0} - \omega_{-0} \big) \Big], \tag{220a}
$$

$$
A^{+\times}(\psi, \omega) = \frac{1}{4h_x h_y} \Big[\psi_{+0} \big(\omega_{+-} - \omega_{++} \big) - \psi_{-0} \big(\omega_{--} - \omega_{-+} \big) - \psi_{0+} \big(\omega_{-+} - \omega_{++} \big) + \psi_{0-} \big(\omega_{--} - \omega_{+-} \big) \Big], \tag{220b}
$$

$$
A^{\times+}(\psi, \omega) = \frac{1}{4h_x h_y} \Big[\psi_{++} \big(\omega_{+0} - \omega_{0+} \big) - \psi_{--} \big(\omega_{0-} - \omega_{-0} \big) - \psi_{-+} \big(\omega_{-0} - \omega_{0+} \big) + \psi_{+-} \big(\omega_{0-} - \omega_{+0} \big) \Big]. \tag{220c}
$$

The indices $(-, 0, +)$ indicate the increment of the corresponding index on the grid relative to the point where the bracket is computed. The integrator for the Poisson equation is

$$
\big\langle \Delta_x \psi_{i,j,k} \big\rangle_t + \big\langle \Delta_y \psi_{i,j,k} \big\rangle_t = \big\langle \omega_{i,j,k} \big\rangle_t . \tag{221}
$$

Here, Δ_x and Δ_y denote the standard centred finite difference second-order derivative with respect to x and y, i.e.,

$$
\Delta_x \psi_{i,j,k} = \frac{\psi_{i,j-1,k} - 2\psi_{i,j,k} + \psi_{i,j+1,k}}{h_x^2},
$$

$$
\Delta_y \psi_{i,j,k} = \frac{\psi_{i,j,k-1} - 2\psi_{i,j,k} + \psi_{i,j,k+1}}{h_y^2},
$$

and the angle brackets $\langle \cdot \rangle_t$ denote an average in time of the form $\frac{1}{4}\begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$, namely,

$$\langle \omega_{i,j,k} \rangle_t = \frac{\omega_{i-1,j,k} + 2\omega_{i,j,k} + \omega_{i+1,j,k}}{4}.$$

It is remarkable that Arakawa's discretisation of the Poisson bracket is recovered. Indeed, a similar derivation was proposed by Salmon and Talley [55]. Our approach, however, is fully covariant, leading to a complete spacetime discretisation, that is a combination of Arakawa's scheme in space with a symplectic integrator in time. Whereas the Arakawa bracket alone guarantees energy conservation only for the spatial discretisation, i.e., up to errors due to the discretisation of the time derivative, the variational integrator is exactly energy preserving.

4.3.4. Simplifications

The integrator (218) for the vorticity equation has one drawback. Even though we are considering a partial differential equation that is first order in time, we need to initialise the fields at two consecutive points in the discrete time domain, that is (218) is a multistep integrator. Nevertheless, we can reduce it to a single-step integrator. Observe that we can rewrite (218) in the following way,

$$0 = \frac{\psi_{i+1} - \psi_i}{2h_t} + \frac{1}{8}\Big[A(\psi_{i+1}, \omega_{i+1}) + A(\psi_{i+1}, \omega_i) + A(\psi_i, \omega_{i+1}) + A(\psi_i, \omega_i) \Big]$$
$$+ \frac{\psi_i - \psi_{i-1}}{2h_t} + \frac{1}{8}\Big[A(\psi_i, \omega_i) + A(\psi_i, \omega_{i-1}) + A(\psi_{i-1}, \omega_i) + A(\psi_{i-1}, \omega_{i-1}) \Big], \tag{222}$$

which is the average over two points in time of the single-step integrator

$$0 = \frac{\psi_i - \psi_{i-1}}{h_t} + \frac{1}{4}\Big[A(\psi_i, \omega_i) + A(\psi_i, \omega_{i-1}) + A(\psi_{i-1}, \omega_i) + A(\psi_{i-1}, \omega_{i-1}) \Big]. \tag{223}$$

Analogously, the Poisson equation (221) is the average over three points in time of

$$\Delta_x \psi_{i,j,k} + \Delta_y \psi_{i,j,k} = \omega_{i,j,k}. \tag{224}$$

The solution of the single-step integrator (223)–(224) is fully determined by specifying the vorticity $\omega = \omega_0$ at $t = t_0$ on the spatial grid. Then ψ_0 can be obtained by (224) and the vorticity can be advected by solving (223). We observe that if the sequence (ω_i, ψ_i) solves (223)–(224) with initial conditions ω_0, then it is also a solution of (218)–(221) with initialisation $(\omega_0, \psi_0, \omega_1, \psi_1)$. Vice versa, if the sequence (ω_i, ψ_i) is a solution of (223)–(224) initialised by using (223)–(224) with the initial data ω_0, then it is also a solution of (223)–(224). Eqs. (223)–(224) are called the underlying one-step method of (218)–(221).

The simplified equations (223) and (224) can also be obtained directly from a discrete Lagrangian as is shown in [22]. This is of course preferable as it allows us to check for symmetries and to compute the discrete conserved quantities more easily. As solutions of the simplified integrator (223)–(224) will also be solutions of (218)–(221), they will satisfy the same conservation laws. Nevertheless, it is possible that (223)–(224) will have additional symmetries and therefore additional invariants.

4.3.5. Discrete conservation laws

The discrete conservation laws of the variational integrator (218) for the vorticity equation are computed in the same way as for the linear advection equation (Section 4.2) except that we have to account for discrete divergence symmetries as introduced in Section 2.6 and that we are using the generators discussed in Section 4.3.2. As in the continuous case, we restrict our analysis to the linear case, that is the Lagrangian (215) discretised according to (216).

Because the vorticity equation is defined on three-dimensional spacetime, the actual computations are tediously long and hardly possible to carry out by hand. We therefore provide only the results, obtained via computer aided calculations, namely

(a) circulation ("mass")

$$h_x h_y \sum_{j,k} \omega_{j,k} = \text{const.}, \tag{225a}$$

(b) enstrophy (L^2 norm)

$$h_x h_y \sum_{j,k} \omega_{j,k}^2 = \text{const.}, \tag{225b}$$

(c) and kinetic energy

$$\frac{h_x h_y}{2} \sum_{j,k} \omega_{j,k} \phi_{j,k} = \text{const.} \tag{225c}$$

It is well known that these quantities are preserved by Arakawa's discrete bracket [54]. In numerical simulations we indeed find that our integrator conserves them to machine accuracy or at least the tolerance of the nonlinear solver, not only in the linear case (when ψ is constant in time), but even in the nonlinear case (with self-consistent streaming function ψ).

4.3.6. Numerical examples

We implemented the simplified variational integrator (223)–(224) using Python [48,49], Cython [56,57], PETSc [58,59] and petsc4py [60]. Visualisation was done using NumPy [50], SciPy [51] and matplotlib [52,53]. The nonlinear system is solved via Picard iteration. Within each nonlinear step the two linear systems corresponding to the vorticity equation and the Poisson equation are solved
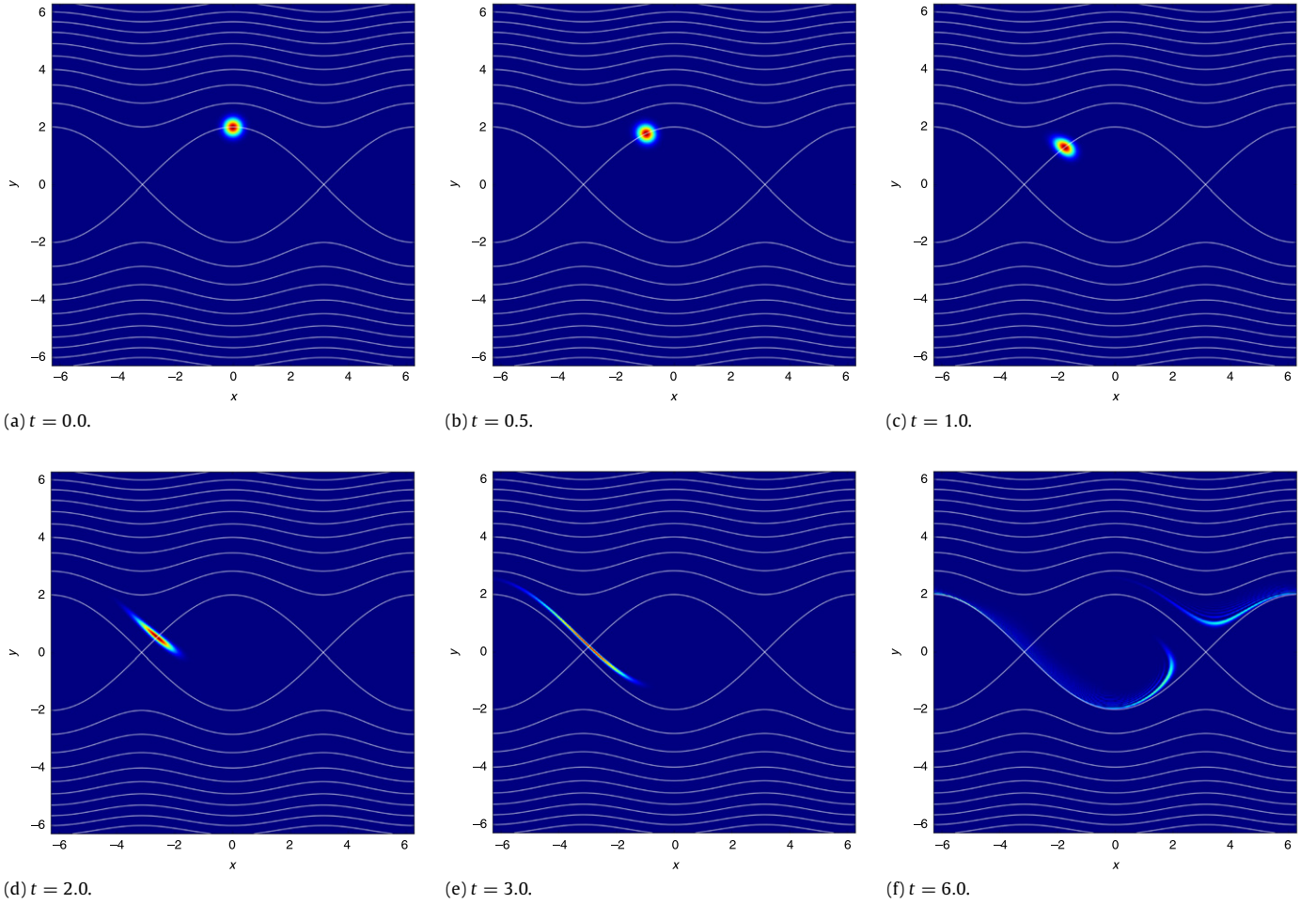
Fig. 9. Linear case with a separatrix. Vorticity ω at $t = 0.0$, 0.5, 1.0, 2.0, 3.0 and 6.0. Contours of the streaming function ψ are also shown.

separately. The vorticity equation is solved with GMRES and the Poisson equation via LU decomposition with SuperLU [61,62]. The tolerance of the nonlinear solver is set to 10^{-10} or smaller, which is reached after $5 - 10$ iterations. Usually, the GMRES solver needs between 5 and 25 iterations to converge with a relative tolerance of 10^{-8} or an absolute tolerance of 10^{-15}.

The linear case

At first we consider the linear case, where we prescribe the streaming function ψ and only solve the vorticity equation. The streaming potential is set to

$$\psi(x, y) = \tfrac{1}{2}y^2 + 1 - \cos(x), \tag{226}$$

while the vorticity is initialised with a localised, symmetric Gaussian,

$$\omega_0(x, y) = f(x, x_0, \sigma_x) f(y, y_0, \sigma_y), \tag{227}$$

with

$$f(z, z_0, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{1}{2}\left[\frac{z - z_0}{\sigma}\right]^2\right), \quad z \in \mathbb{R}. \tag{228}$$

The domain is $\Omega = [-2\pi, +2\pi) \times [-2\pi, +2\pi)$ with periodic boundaries, the spatial resolution is 1024×1024 and the timestep is 10^{-2}.

The parameters are set to $x_0 = 0$, $y_0 = 2$ and $\sigma_x = \sigma_y = 0.2$, so that the Gaussian is placed on the separatrix of the streaming function (c.f., Fig. 9). As the Gaussian is moving along the separatrix, it is stretched while the contours of the vorticity are preserved. In fact, if a field ω is advected by a smooth flow, the topology of the contours of ω should be preserved. This behaviour appears to be respected by the integrator. In Fig. 10(a), the contours for different values of the vorticity are shown. Until $t = 3$ all three contours stay intact. At about $t = 6$ the stretching of the contours is so strong that resolution becomes insufficient and only the outermost contour is still preserved. Towards the centre and the boundaries in x the effect of dispersion becomes visible. Fig. 10(a) shows the evolution of the errors of circulation, enstrophy and energy.
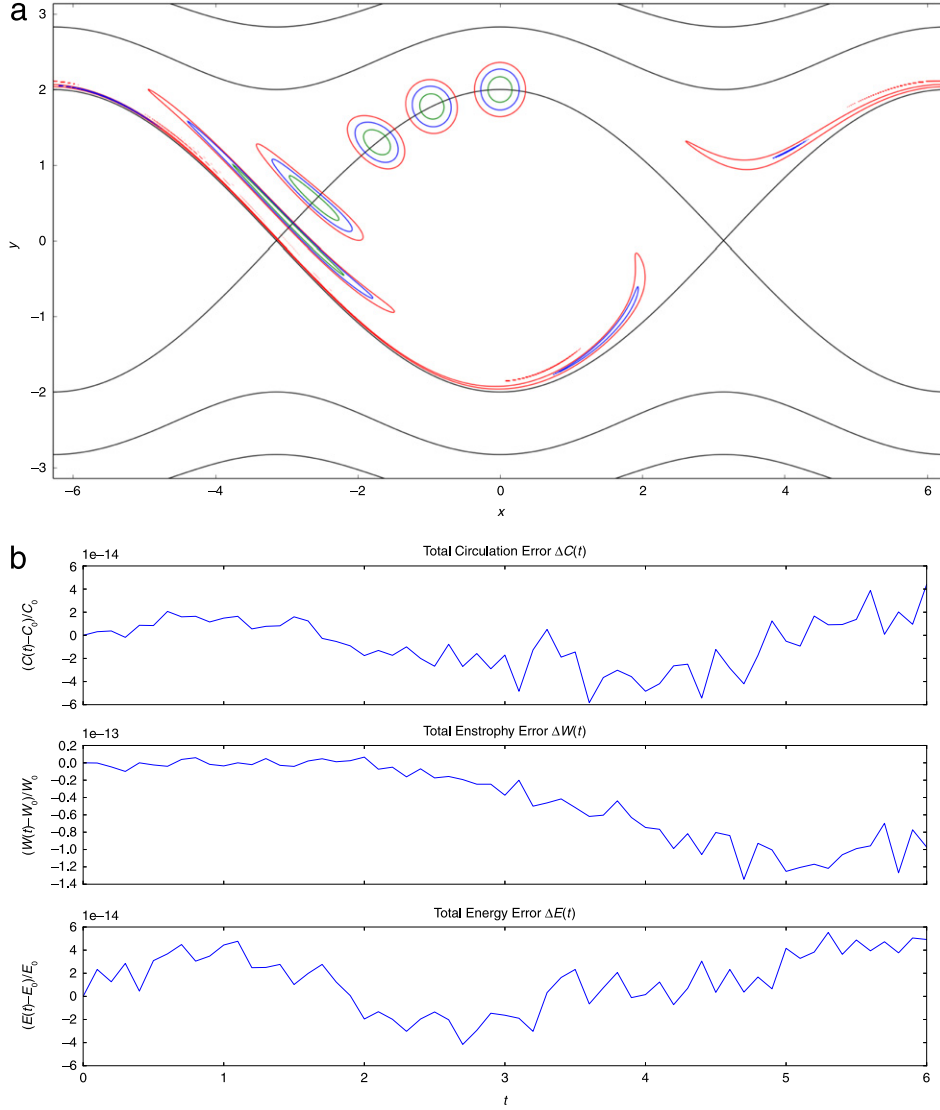
Fig. 10. Linear case with a separatrix. (a) Contours of the streaming function ψ (black) and of the vorticity ω (at $0.2 \max \omega_0$, $0.4 \max \omega_0$ and $0.7 \max \omega_0$) at times $t = 0.0$, 0.5, 1.0, 2.0, 3.0 and 6.0. (b) Conservation laws.

Lamb dipole

The lamb dipole [63] is a stationary solution of the vorticity equation which is at rest in its frame of reference. The vorticity is initialised as

$$
\omega_0(x, y) = \begin{cases} 2\lambda U \cos\theta \dfrac{J_1(\lambda r)}{J_0(\lambda R)} & r \le R, \\ 0 & r > R, \end{cases}
\tag{229}
$$

which leads to a dipole of radius R moving in the y-direction with velocity U. Here, r and θ denote polar coordinates, i.e.,

$$
r = \sqrt{x^2 + y^2} \quad \text{and} \quad \theta = \arctan(y/x),
\tag{230}
$$

J_i is the ith Bessel function of first kind and λR is the first root of J_1. In the example, we use $R = 0.2$ and $U = 1$. The domain is $\Omega = [-1, +1) \times [-1, +1)$ with periodic boundaries, the spatial resolution is 1024×1024 and the timestep is 10^{-3}. The tolerance of the nonlinear solver is set to 10^{-12}.

After 10,000 timesteps (at $t = 10$) no distortion of the lamb dipole is visible (compare Figs. 11(a) and (b)). The dipole keeps its shape just as it is supposed to. It does not end up exactly at the initial position due to the finite size of the box which reduces the velocity of the reference frame [63, section IV.B]. Fig. 11(c) shows that circulation, enstrophy and energy are preserved to about machine accuracy. There is a small drift in the errors of enstrophy and energy, but this is of $\mathcal{O}(10^{-13})$ and therefore comparable to the tolerance of the nonlinear solver.

Gaussian vortex

Vorticity is initialised by a nonsymmetric Gaussian,

$$
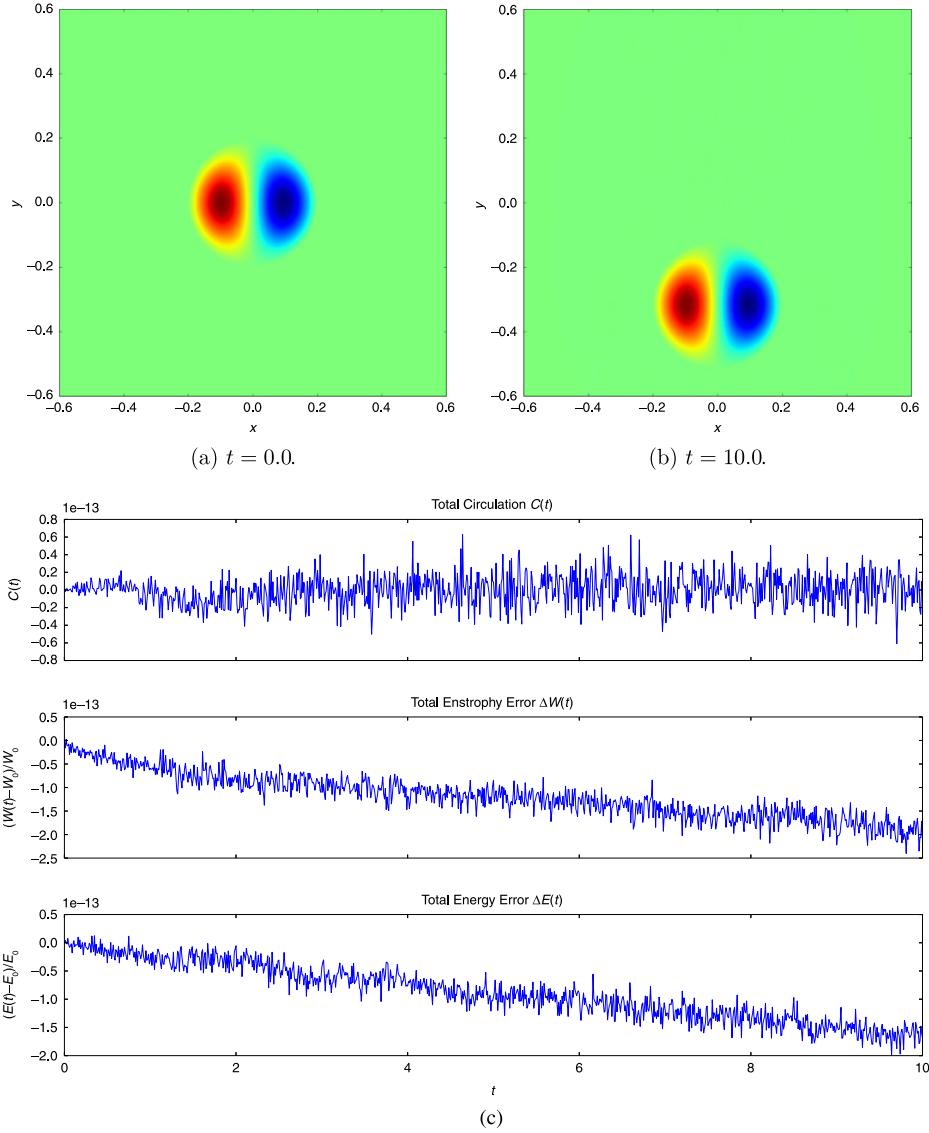\omega_0(x, y) = f(x, x_0, \sigma_x) f(y, y_0, \sigma_y),
\tag{231}
$$

Fig. 11. Lamb Dipole. Top: Vorticity ω at $t = 0.0$ and 10.0. Bottom: Conservation Laws.

with

$$f(x, x_0, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{1}{2}\left[\frac{x - x_0}{\sigma}\right]^2\right), \tag{232}$$

which develops into a spiral structure. The domain is $\Omega = [-1, +1) \times [-1, +1)$ with periodic boundaries, the spatial resolution is 1024×1024 and the timestep is 10^{-2}. The tolerance of the nonlinear solver is 10^{-11}. We set $\sigma_x = 0.1$, $\sigma_y = 0.2$ and $x_0 = y_0 = 0$. The solution follows the expected dynamics (c.f., Fig. 12) and conservation of circulation, enstrophy and energy is respected to machine accuracy (c.f., Fig. 14(a)).

Vortex sheet rollup

The vortex sheet is a popular model in fluid dynamics used to approximate shear layers. Following [64], we use initial conditions

$$u_0 = \begin{pmatrix} \tanh(\rho(y - 0.25)) & \text{for } y \leq 0.5, \\ \tanh(\rho(0.75 - y)) & \text{for } y > 0.5, \end{pmatrix}\quad 0.05 \sin(2\pi x) \end{pmatrix}^T, \tag{233}$$

which corresponds to the initial vorticity

$$\omega_0(x, y) = \nabla_\perp \cdot u_0 = 0.1\pi \cos(2\pi x) + \begin{cases} -\dfrac{\rho}{\cosh^2(\rho(y - 0.25))} & y \leq r, \\[2ex] +\dfrac{\rho}{\cosh^2(\rho(0.75 - y))} & y > r. \end{cases} \tag{234}$$

The domain is $\Omega = [0, 1) \times [0, 1)$ with periodic boundaries, discretised by a grid of 1024×1024 points, and the timestep is $h_t = 10^{-3}$. The tolerance of the nonlinear solver is 10^{-10} and the parameter $\rho = 30$. We observe good qualitative agreement of the solution in Fig. 13 with [64, p. 328] together with excellent conservation properties (c.f., Fig. 14(b)).
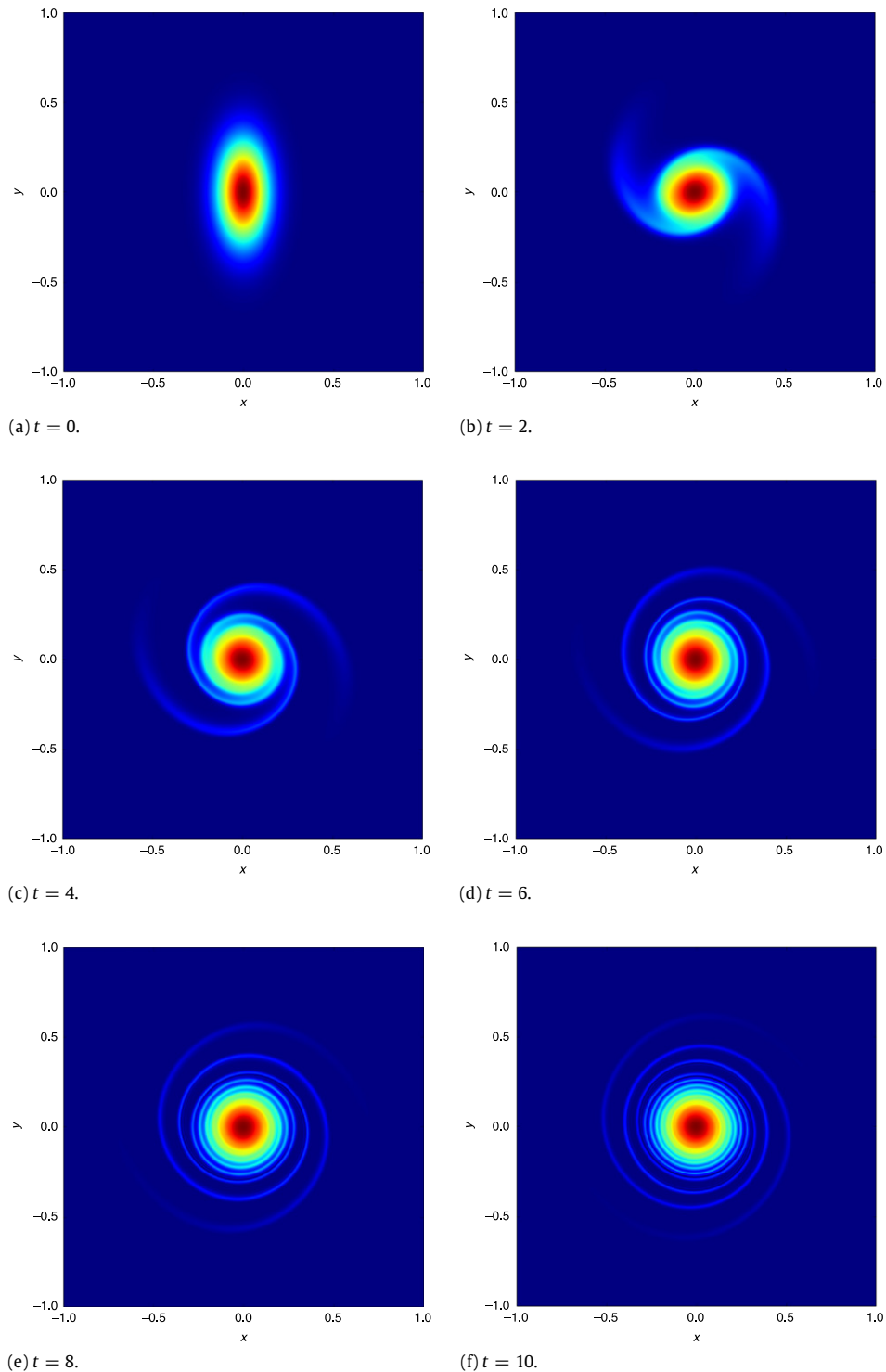
(a) $t = 0$.

(b) $t = 2$.

(c) $t = 4$.

(d) $t = 6$.

(e) $t = 8$.

(f) $t = 10$.

Fig. 12. Gaussian Vortex. Vorticity ω at $t = 0, 2, 4, 6, 8, 10$.

5. Summary and outlook

We created a link between variational integrators, formal Lagrangians and Ibragimov's theory of conservation laws for arbitrary differential equations. The proposed method allows us to derive variational discretisation schemes for potentially arbitrary systems of differential equations, even the ones that do not possess a classical variational formulation. Thereby we were able to extend the applicability of the variational integrator method to a class of systems much larger than originally envisaged. The main strength of our method is that it allows for the straight forward design of numerical schemes that respect certain conservation laws of the system at hand.

We presented an introduction to the theory of variational integrators that tries to make the geometric framework of the variational integrators available also to non specialists. Thereby we hope to make this class of methods accessible to a wider audience. We extended the discrete Noether theorem to include discrete divergence symmetries.
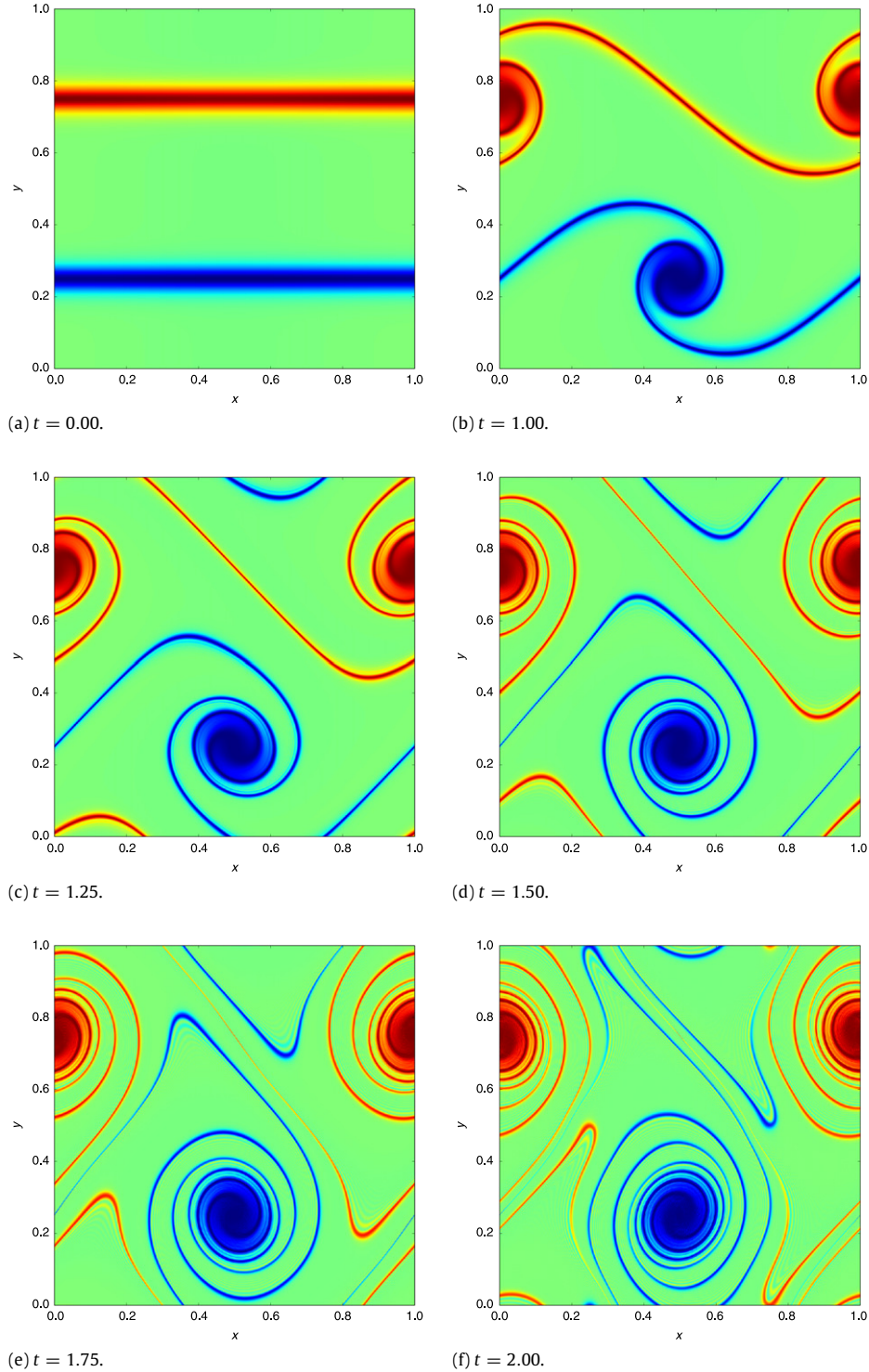
(a) $t = 0.00$.

(b) $t = 1.00$.

(c) $t = 1.25$.

(d) $t = 1.50$.

(e) $t = 1.75$.

(f) $t = 2.00$.

Fig. 13. Vortex sheet rollup. Vorticity ω at $t = 0.00, 1.00, 1.25, 1.50, 1.75, 2.00$.

The power of the method was demonstrated by several examples that are prototypical for problems arising in fluid dynamics and plasma physics. We emphasised the analysis of discrete conservation laws, which is seldom found in the variational integrator literature, and verified these theoretical properties in numerical experiments. In follow-up papers we will present numerical results for the Vlasov–Poisson system [22] as well as ideal and reduced magnetohydrodynamics [23,24]. There, explicit computations will demonstrate the favourable properties of the variational integrators for more elaborate and challenging applications.

Remarkably, we recovered Arakawa's discretisation of the Poisson bracket combined with a symplectic integrator in time. That is, we constructed a spacetime generalisation of Arakawa's discretisation. With our method, it is also straight-forward to generalise Arakawa's method to higher spatial dimensions and to higher order schemes.

An open question within our framework is the meaning of the discrete multisymplectic form arising from the boundary terms of the action sum (see [6] for further details) and its restriction to the original system. We observed that the multisymplectic form vanishes
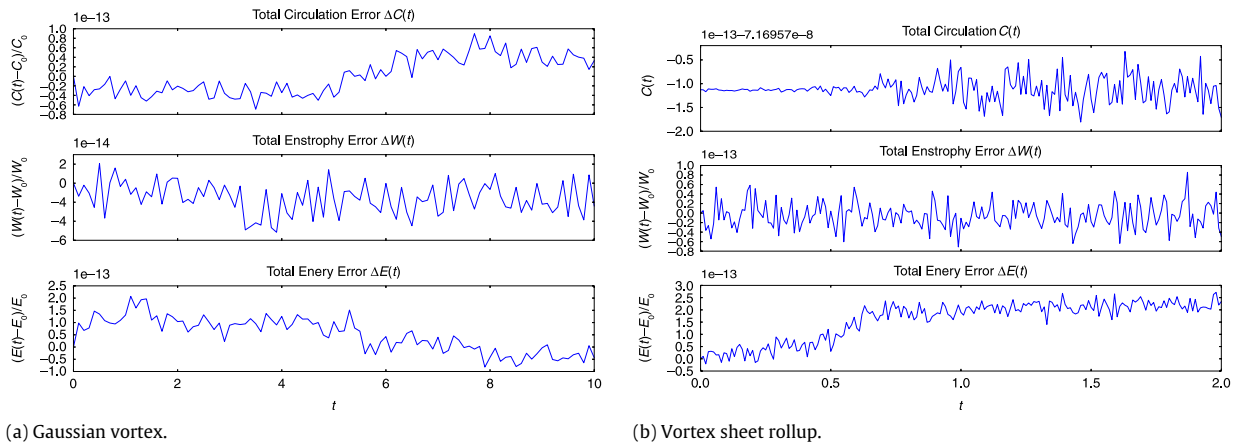
(a) Gaussian vortex.

(b) Vortex sheet rollup.

Fig. 14. Conservation Laws. The errors in the circulation, enstrophy and energy are all of the order of the machine accuracy.

identically if the extended system is self-adjoint in the sense of Ibragimov. This is a topic we have not included in our discussion but one that certainly deserves attention.

The limits of the finite-difference approach to the discretisation of the Lagrangian became obvious in several places in this work. The treatment of discrete divergence symmetries is not straight-forward so that we are limited to global conservation laws. We cannot treat horizontal transformations in the symmetry generator but only vertical transformations. This is a problem for more complicated systems like the vorticity equations or the nonlinear advection equation (inviscid Burgers equation). For nonlinear systems there is the additional complication that even if the system is self-adjoint on the continuous level, it will in general not be self-adjoint on the discrete level due to the absence of a discrete Leibniz rule. Moreover, we are limited to low-order schemes. While it is in principle possible to design higher-order methods, this quickly becomes very cumbersome and confusing, especially in the analysis of the discrete conservation laws. All those issues appear easier to deal with in a Galerkin framework with finite elements or splines as basis functions. The former has already been considered to a certain extend [45], that is for the discretisation of the Lagrangian and the approximation of the action integral, but not for the analysis of conservation laws. The latter is a topic under active development [46].

So far, we only considered examples of advection–diffusion type. A detailed analysis of the applicability of the proposed method to other types of equations would be most interesting.

Acknowledgements

References

[1] Snorre H. Christiansen, Hans Z. Munthe-Kaas, Brynjulf Owren, Topics in structure-preserving discretization, Acta Numer. 20 (2011) 1–119. http://dx.doi.org/10.1017/S096249291100002X.
[2] Ernst Hairer, Christian Lubich, Gerhard Wanner, Geometric Numerical Integration, Springer, 2006.
[3] Chris J. Budd, Matthew D. Piggott, Geometric integration and its applications, in: Handbook of numerical analysis, North-Holland, 2000, pp. 35–139.
[4] Jeffrey M. Wendlandt, Jerrold E. Marsden, Mechanical integrators derived from a discrete variational principle, Physica D 106 (1997) 223–246. http://dx.doi.org/10.1016/S0167-2789(97)00051-1.
[5] Jerrold E. Marsden, Jeffrey M. Wendlandt, Mechanical systems with symmetry, variational principles, and integration algorithms, in: Mark Alber, Bei Hu, Joachim Rosenthal (Eds.), Current and Future Directions in Applied Mathematics, Birkhäuser, Boston, 1997, pp. 219–261. http://dx.doi.org/10.1007/978-1-4612-2012-1. ISBN 978-1-4612-7380-6.
[6] Jerrold E. Marsden, George W. Patrick, Steve Shkoller, Multisymplectic geometry, variational integrators, and nonlinear PDEs, Comm. Math. Phys. 199 (1998) 351–395. http://dx.doi.org/10.1007/s002200050505. arXiv:math/9807080.
[7] Shinar O. Kouranbaeva, Steve Shkoller, A variational approach to second-order multisymplectic field theory, J. Geom. Phys. 25 (2000) 333–366. http://dx.doi.org/10.1016/S0393-0440(00)00012-7. arXiv:math/9909100.
[8] Jerrold E. Marsden, Matthew West, Discrete mechanics and variational integrators, Acta Numer. 10 (2001) 357–514. http://dx.doi.org/10.1017/S096249290100006X.
[9] Adrian Lew, Jerrold E. Marsden, Michael Ortiz, Matthew West, An Overview of Variational Integrators, volume Finite Element Methods: 1970's and Beyond. Theory and engineering applications of computational methods, pages 98–115. International Center for Numerical Methods in Engineering (CIMNE), Barcelona, 2004b.
[10] A. Lew, J.E. Marsden, M. Ortiz, M. West, Variational time integrators, Internat. J. Numer. Methods Engrg. 60 (1) (2004) 153–212. http://dx.doi.org/10.1002/nme.958.
[11] Jorge V. Jose, Eugene J. Saletan, Classical Dynamics: A Contemporary Approach, Cambridge University Press, 1998.
[12] Vladimir I. Arnold, Mathematical Methods of Classical Mechanics, Springer, 1989.
[13] Darryl D. Holm, Tanya Schmah, Cristina Stoica, Geometric Mechanics and Symmetry, Oxford University Press, 2009.
[14] Jerrold E. Marsden, Tudor S. Ratiu, Introduction to Mechanics and Symmetry, Springer, 2002.
[15] Ralph Abraham, Jerrold E. Marsden, Foundations of Mechanics, American Mathematical Society, 1978. URL http://authors.library.caltech.edu/25029/.
[16] Emmy Noether, Invariante variationsprobleme, Nachrichten der Königlichen Gesellschaft der Wissenschaften zu Göttingen (1918) 235–257. arXiv:physics/0503066.
[17] Yvette Kosmann-Schwarzbach, The Noether Theorems: Invariance and Conservation Laws in the Twentieth Century, Springer, 2010.
[18] Robert W. Atherton, George M. Homsy, On the existence and formulation of variational principles for nonlinear differential equations, Stud. Appl. Math. 54 (1975) 31–60.
[19] Nail H. Ibragimov, Integrating factors, adjoint equations and lagrangians, J. Math. Anal. Appl. 318 (2006) 742–757. http://dx.doi.org/10.1016/j.jmaa.2005.11.012.
[20] Nail H. Ibragimov, A new conservation theorem, J. Math. Anal. Appl. 333 (2007) 311–328. http://dx.doi.org/10.1016/j.jmaa.2006.10.078.
[21] Nail H. Ibragimov, Quasi-self-adjoint differential equations, Arch. ALGA 4 (2007) 55–60.

[22] Michael Kraus, Omar Maj, Eric Sonnendrücker, Variational Integrators for the Vlasov-Poisson System, 2015, in preparation.
[23] Michael Kraus, Omar Maj, Variational Integrators for Ideal Magnetohydrodynamics, 2015, in preparation.
[24] Michael Kraus, Emanuele Tassi, Variational Integrators for Reduced Magnetohydrodynamics, 2015, in preparation.
[25] David J. Saunders, The Geometry of Jet Bundles, Cambridge University Press, 1989.
[26] Ivan Kolar, Jan Slovak, Peter W. Michor, Natural Operations in Differential Geometry, Springer, 1993. URL http://www.emis.de/monographs/KSM/.
[27] Peter J. Olver, Equivalence, Invariants and Symmetry, Cambridge University Press, 1995.
[28] Vladimir G. Ivancevic, Tijana T. Ivancevic, Applied Differential Geometry, World Scientific, 2007.
[29] Cédric M. Campos, Geometric methods in classical field theory and continuous media (Ph.D. thesis), Universidad Autónoma de Madrid, 2010.
[30] Mark J. Gotay, James Isenberg, Jerrold E. Marsden, Momentum maps and classical fields, 1998.
[31] Jerrold E. Marsden, Sergey Pekarsky, Steve Shkoller, Matthew West, Variational methods, multisymplectic geometry and continuum mechanics, J. Geom. Phys. 38 (2001) 253–284. http://dx.doi.org/10.1016/S0393-0440(00)00066-8. arXiv:math/0005034.
[32] Adrian Lew, Jerrold E. Marsden, Michael Ortiz, Matthew West, Asynchronous variational integrators, Arch. Ration. Mech. Anal. 167 (2003) 85–146. http://dx.doi.org/10.1007/s00205-002-0212-y.
[33] Peter J. Olver, Applications of Lie Groups to Differential Equations, Springer, 1993.
[34] Alexander P. Veselov, Integrable discrete-time systems and difference operators, Funct. Anal. Appl. 22 (1988) 83–93. http://dx.doi.org/10.1007/BF01077598.
[35] Alexander P. Veselov, Integrable lagrangian correspondences and the factorization of matrix polynomials, Funct. Anal. Appl. 25 (1991) 112–122. http://dx.doi.org/10.1007/BF01079590.
[36] Jürgen Moser, Alexander P Veselov, Discrete versions of some classical integrable systems and factorization of matrix polynomials, Comm. Math. Phys. 139 (1991) 217–243.
[37] Philip J. Morrison, Hamiltonian description of the ideal fluid, Rev. Modern Phys. 70 (1998) 467–521. http://dx.doi.org/10.1103/RevModPhys.70.467.
[38] Ralph Abraham, Jerrold E. Marsden, Tudor S. Ratiu, Manifolds, Tensor Analysis and Application, Springer, 1988.
[39] F. Bampi, A. Morro, The inverse problem of the calculus of variations applied to continuum physics, J. Math. Phys. 23 (1982) 2312–2321. http://dx.doi.org/10.1063/1.525322.
[40] Enzo Tonti, Variational formulations of nonlinear differential equations, Bulletins de l'Academie Royale des Sciences, des Lettres et des Beaux Arts de Belgique LV (1969) 137–165, 262–278.
[41] R.L. Seliger, G.B. Whitham, Variational principles in continuum mechanics, Proc. R. Soc. A 305 (1968) 1–25. http://dx.doi.org/10.1098/rspa.1968.0103.
[42] Dmitry Pavlov, Patrick Mullen, Yiying Tong, Eva Kanso, Jerrold E. Marsden, Mathieu Desbrun, Structure-preserving discretization of incompressible fluids, Physica D 240 (2011) 443–458. http://dx.doi.org/10.1016/j.physd.2010.10.012. arXiv:0912.3989.
[43] Evan S. Gawlik, Patrick Mullen, Dmitry Pavlov, Jerrold E. Marsden, Mathieu Desbrun, Geometric, variational discretization of continuum theories, Physica D 240 (2011) 1724–1760. arXiv:1010.4851 http://dx.doi.org/10.1016/j.physd.2011.07.011.
[44] Nail H. Ibragimov, Mariano Torrisi, Rita Tracinà, Quasi self-adjoint nonlinear wave equations, J. Phys. A 43 (2010). http://dx.doi.org/10.1088/1751-8113/43/44/442001.
[45] Jing-Bo Chen, Variational integrators and the finite element method, Appl. Math. Comput. 196 (2008) 941–958.
[46] Michael Kraus, Spline Variational Integrators, 2015, in preparation.
[47] Uri M. Ascher, Robert I. McLachlan, Multisymplectic box schemes and the korteweg–de vries equation, Appl. Numer. Math. 48 (2004) 255–269. http://dx.doi.org/10.1016/j.apnum.2003.09.002.
[48] The Python Programming Language. http://www.python.org/.
[49] Hans Petter Langtangen, A Primer on Scientific Programming with Python, fourth ed., Springer, 2014.
[50] S. van der Walt, S.C. Colbert, G. Varoquaux, The numpy array: A structure for efficient numerical computation, Comput. Sci. Eng. 13 (2011) 22–30. http://dx.doi.org/10.1109/MCSE.2011.37. arXiv:1102.1523.
[51] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python. http://www.scipy.org/.
[52] The matplotlib Plotting Library. http://www.matplotlib.org/.
[53] J.D. Hunter, Matplotlib: A 2d graphics environment, Comput. Sci. Eng. 9 (2007) 90–95.
[54] Akio Arakawa, Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow. Part i, J. Comput. Phys. 1 (1966) 119–143. http://dx.doi.org/10.1016/0021-9991(66)90015-5.
[55] Rick Salmon, Lynne D. Talley, Generalizations of arakawa's jacobian, J. Comput. Phys. 83 (1989) 247–259. http://dx.doi.org/10.1016/0021-9991(89)90118-6.
[56] R. Bradshaw, S. Behnel, D.S. Seljebotn, G. Ewing, et al. The Cython Compiler. http://www.cython.org/.
[57] S. Behnel, R. Bradshaw, C. Citro, L. Dalcin, D.S. Seljebotn, K. Smith, Cython: The best of both worlds, Comput. Sci. Eng. 13 (2011) 31–39. http://dx.doi.org/10.1109/MCSE.2010.118.
[58] Satish Balay, Shrirang Abhyankar, Mark F. Adams, Jed Brown, Peter Brune, Kris Buschelman, Victor Eijkhout, William D. Gropp, Dinesh Kaushik, Matthew G. Knepley, Lois Curfman McInnes, Karl Rupp, Barry F. Smith, Hong Zhang, PETSc Web page. http://www.mcs.anl.gov/petsc.
[59] Satish Balay, Shrirang Abhyankar, Mark F. Adams, Jed Brown, Peter Brune, Kris Buschelman, Victor Eijkhout, William D. Gropp, Dinesh Kaushik, Matthew G. Knepley, Lois Curfman McInnes, Karl Rupp, Barry F. Smith, Hong Zhang, PETSc users manual. Technical Report ANL-95/11 - Revision 3.5, Argonne National Laboratory, 2014. URL http://www.mcs.anl.gov/petsc.
[60] L. Dalcin, P. Kler, R. Paz, A. Cosimo, Parallel distributed computing using python, Adv. Water Resour. 34 (2011) 1124–1139. http://dx.doi.org/10.1016/j.advwatres.2011.04.013.
[61] X.S. Li, J.W. Demmel, J.R. Gilbert, iL. Grigori, M. Shao, I. Yamazaki, SuperLU Users' Guide. Technical Report LBNL-44289, Lawrence Berkeley National Laboratory, 1999.
[62] Xiaoye S. Li, An overview of SuperLU: algorithms, implementation, and user interface, ACM Trans. Math. Softw. 31 (2005) 302–325.
[63] A.H. Nielsen, J. Juul Rasmussen, Formation and temporal evolution of the lamb-dipole, Phys. Fluids 9 (1997) 982–991. http://dx.doi.org/10.1063/1.869193.
[64] M.O. Deville, P.F. Fischer, E.H. Mund, High-Order Methods for Incompressible Fluid Flow, Cambridge University Press, 2002.