FOUNDATIONS OF
COMPUTATIONAL
MATHEMATICS
The Journal of the Society for the Foundations of Computational Mathematics

Functional Equivariance and Conservation Laws in Numerical Integration

Robert I. McLachlan[1] · Ari Stern[2]

© SFoCM 2022

Abstract
Preservation of linear and quadratic invariants by numerical integrators has been well studied. However, many systems have linear or quadratic observables that are not invariant, but which satisfy evolution equations expressing important properties of the system. For example, a time-evolution PDE may have an observable that satisfies a local conservation law, such as the multisymplectic conservation law for Hamiltonian PDEs. We introduce the concept of *functional equivariance*, a natural sense in which a numerical integrator may preserve the dynamics satisfied by certain classes of observables, whether or not they are invariant. After developing the general framework, we use it to obtain results on methods preserving local conservation laws in PDEs. In particular, integrators preserving quadratic invariants also preserve local conservation laws for quadratic observables, and symplectic integrators are multisymplectic.

Communicated by Arieh Iserles.

✉ Ari Stern
  stern@wustl.edu

  Robert I. McLachlan
  r.mclachlan@massey.ac.nz

1   School of Fundamental Sciences, Massey University, Palmerston North, New Zealand

2   Department of Mathematics and Statistics, Washington University in St. Louis, St. Louis, USA

 Springer

1 Introduction

In numerical ordinary differential equations (ODEs), it is known that all B-series methods (including Runge–Kutta methods) preserve linear invariants, while only certain ones preserve quadratic invariants. Linear invariants arising in physical systems include mass, charge, and linear momentum; quadratic invariants include angular momentum and other momentum maps, as well as the canonical symplectic form for Hamiltonian systems. See Hairer, Lubich, and Wanner [15] and references therein.

However, for partial differential equations (PDEs) describing time evolution, it is desirable for a numerical integrator to preserve not only global invariants but also local conservation laws. For instance, the evolution may preserve total mass (a global invariant), but the mass in a particular region may change by flowing through the boundary of the region (a local conservation law). Another example is the canonical multisymplectic conservation law for Hamiltonian PDEs, which is a quadratic local conservation law for the variational equation. Focusing only on global invariants overlooks this more granular, local form of conservativity.

This paper develops a new framework for the preservation of such properties by numerical integrators. We do so by answering a much more general question: When does a numerical integrator preserve the evolution of certain classes of observables (e.g., linear, quadratic), even when those observables are not invariants? This includes not only global invariants, as previously studied, but also local conservation laws and other balance laws encountered in both conservative and dissipative dynamical systems.

The main idea of our approach is summarized as follows. Suppose $y = y(t)$ evolves in a (finite- or infinite-dimensional) Banach space Y according to $\dot{y} = f(y)$. Given a functional $F \in C^1(Y)$, the chain rule implies that $z = F(y)$ evolves according to $\dot{z} = F'(y)f(y)$. Now, if Φ is a numerical integrator, let $\Phi_f : y_0 \mapsto y_1$ denote its application to the original system $\dot{y} = f(y)$, and let $\Phi_g : (y_0, z_0) \mapsto (y_1, z_1)$ denote its application to the augmented system

$$\dot{y} = f(y), \qquad \dot{z} = F'(y)f(y), \tag{1}$$

corresponding to the vector field $g(y, z) = \big(f(y), F'(y)f(y)\big)$. We say that Φ is F-*functionally equivariant* if Φ_g preserves the relation $z = F(y)$, i.e., $\Phi_g : \big(y_0, F(y_0)\big) \mapsto \big(y_1, F(y_1)\big)$, for all vector fields f on Y. In other words, the following diagram commutes:

$$
\begin{array}{ccc}
y_0 & \xmapsto{\;\;\Phi_f\;\;} & y_1 \\
{\scriptstyle(\mathrm{id},F)}\big\downarrow & & \big\downarrow{\scriptstyle(\mathrm{id},F)} \\
(y_0, z_0) & \xmapsto{\;\;\Phi_g\;\;} & (y_1, z_1)\,.
\end{array}
$$

This is weaker than equivariance in the usual sense, since the diagram need only commute for (1), not arbitrary (id, F)-related vector fields. Preserving invariants becomes the special case where the augmented equation reads $\dot{z} = 0$ and the integrator leaves z constant.

We develop a theory of functional equivariance and show that it provides a useful tool kit for understanding the behavior of (especially affine and quadratic) observables, including local conservation laws and multisymplecticity. The paper is organized as follows:

- Section 2 characterizes the functional equivariance of a large class of numerical integrators, including B-series methods, and explores some consequences for both conservative and non-conservative dynamical systems. The main result, Theorem 2.9, shows that a method is functionally equivariant for a class of observables if and only if it preserves invariants in that class. In particular, all B-series methods are affine functionally equivariant, and those preserving quadratic invariants are quadratic functionally equivariant.
- Section 3 applies this framework to local conservation laws for PDEs and spatially semidiscretized PDEs. In particular, affine/quadratic functionally equivariant numerical integrators are seen to preserve discrete-time local conservation laws for affine/quadratic observables.
- Section 4 applies this framework to the multisymplectic conservation law for canonical Hamiltonian PDEs and spatially semidiscretized PDEs. We show that multisymplectic semidiscretization in space, followed by a symplectic integrator in time, yields a multisymplectic method in spacetime. We also show that hybrid finite elements may be used for multisymplectic semidiscretization, generalizing the results of McLachlan and Stern [22] to time-evolution problems.
- Finally, Sect. 5 extends the results from the class of methods considered in Sect. 2 to additive and partitioned methods, including additive/partitioned Runge–Kutta methods and splitting/composition methods.

We remark that many of the results, particularly in Sects. 2 and 5, are obtained using only the equivariance properties of methods with respect to affine maps, rather than representing them in terms of trees or Runge–Kutta tableaux. In particular, Theorem 2.12 gives a new, tree-free proof that B-series methods are closed under differentiation, while Theorem 5.20 generalizes this to additive and partitioned methods.

2 Functional Equivariance

2.1 Basic Definitions and Results

Let Φ be a one-step numerical integrator, whose application to a vector field $f \in \mathfrak{X}(Y)$ with time-step size Δt gives a map $\Phi_{\Delta t, f} \colon Y \to Y$, $y_0 \mapsto y_1$. All the methods we will consider have $\Phi_{\Delta t, f} = \Phi_{1, \Delta t f}$, so it suffices to consider integrator maps $\Phi_f := \Phi_{1, f}$ with unit time step. When we refer to a numerical integrator, we mean the entire collection of maps $\Phi = \left\{ \Phi_f : f \in \mathfrak{X}(Y), \ Y \text{ a Banach space} \right\}$.[1]

Remark 2.1 While this definition covers a large class of numerical integrators, including B-series methods, other classes of methods require additional data besides f in

---

[1] For some methods, such as implicit Runge–Kutta methods, $\Phi_{\Delta t, f}(y)$ might only be defined for sufficiently small Δt. Including such integrators requires only the minor modification of viewing Φ_f as a partial function.

order to define an integrator map, e.g., an additive decomposition of f or a partitioning of Y. In Sect. 5, we will discuss how the results of this section generalize to such methods, including additive/partitioned Runge–Kutta methods and splitting/composition methods.

The main class of numerical integrators of this type that we will consider are B-series methods, which McLachlan, Modin, Munthe-Kaas, and Verdier [18] proved are completely characterized by the following property of affine equivariance.

Definition 2.2 Given an affine map $A\colon Y \to U$, a pair of vector fields $f \in \mathfrak{X}(Y)$ and $g \in \mathfrak{X}(U)$ is *A-related* if $A' \circ f = g \circ A$. A numerical integrator Φ is *affine equivariant* if $A \circ \Phi_f = \Phi_g \circ A$ for all A-related f and g, all affine maps A, and all Banach spaces Y and U.

Remark 2.3 This is consistent with the definition of affine equivariance in [18]. We distinguish it from the weaker definition in Munthe-Kaas and Verdier [23], where the condition above is required only for affine isomorphisms rather than all affine maps.

Definition 2.4 Given a Gâteaux differentiable map $F\colon Y \to Z$ and $f \in \mathfrak{X}(Y)$, define $g \in \mathfrak{X}(Y \times Z)$ by $g(y, z) = \big(f(y), F'(y)f(y)\big)$. We say that a numerical integrator Φ is *F-functionally equivariant* if $(\mathrm{id}, F) \circ \Phi_f = \Phi_g \circ (\mathrm{id}, F)$ for all $f \in \mathfrak{X}(Y)$. That is, if $\Phi_f\colon y_0 \mapsto y_1$, then $\Phi_g\colon \big(y_0, F(y_0)\big) \mapsto \big(y_1, F(y_1)\big)$. Given a class of maps $\mathcal{F}$, we say that Φ is *$\mathcal{F}$-functionally equivariant* if this holds for all $F \in \mathcal{F}(Y, Z)$ and all Banach spaces Y and Z.

This is a slight generalization of the situation considered in the introduction: Z may now be any Banach space rather than $\mathbb{R}$, and F is only required to be Gâteaux differentiable rather than C^1. Note that $g \in \mathfrak{X}(Y \times Z)$ is precisely the vector field corresponding to the augmented system (1).

Example 2.5 (Runge–Kutta methods) An s-stage Runge–Kutta method has the form

$$Y_i = y_0 + \Delta t \sum_{j=1}^{s} a_{ij} f(Y_j), \qquad i = 1, \ldots, s,$$

$$y_1 = y_0 + \Delta t \sum_{i=1}^{s} b_i f(Y_i),$$

where a_{ij} and b_i are given coefficients defining the method. When this method is applied to the augmented system (1), we augment the method by

$$z_1 = z_0 + \Delta t \sum_{i=1}^{s} b_i F'(Y_i) f(Y_i).$$

Note that the internal stages $Z_1, \ldots, Z_s$ are not needed, since the augmented vector field depends only on y. Hence, for a Runge–Kutta method, F-functional equivariance

says that

$$F(y_1) = F(y_0) + \Delta t \sum_{i=1}^{s} b_i F'(Y_i) f(Y_i).$$

In particular, if F is an invariant of f, so that $F'(y) f(y) = 0$ for all $y \in Y$, then the terms of this sum vanish, and we get $F(y_1) = F(y_0)$.

Proposition 2.6 *Every affine equivariant method is affine functionally equivariant.*

Proof If $F: Y \to Z$ is an affine map, then so is $(\mathrm{id}, F): Y \to Y \times Z$. Since the vector fields f and g in Definition 2.4 are (id, F)-related, the conclusion follows by affine equivariance. $\qquad\square$

Remark 2.7 The converse is generally not true. For instance, the map $y_0 \mapsto y_0 + (f \operatorname{div} f)(y_0)$, which is defined for finite-dimensional Y, is seen to be affine functionally equivariant but is not affine equivariant except in the weaker sense mentioned in Remark 2.3. This is an example of an "aromatic" series that is not a B-series, cf. Munthe-Kaas and Verdier [23].

Since we are concerned for the time being with affine equivariant numerical integrators, it is natural to make the following assumptions on $\mathcal{F}$. This includes cases where $\mathcal{F}$ contains not only affine maps but also quadratic or higher-degree polynomial maps.

Assumption 2.8 The class of maps $\mathcal{F}$ satisfies the following:

- $\mathcal{F}(Y, Y)$ contains the identity map for all Y;
- $\mathcal{F}(Y, Z)$ is a vector space for all Y and Z;
- $\mathcal{F}$ is invariant under composition with affine maps, in the following sense: If $A: Y \to U$ and $B: V \to Z$ are affine and $F \in \mathcal{F}(U, V)$, then $B \circ F \circ A \in \mathcal{F}(Y, Z)$.

As noted in the introduction, preservation of invariants may be seen as a special case of functional equivariance, so one might expect the latter property to be strictly stronger than the former. Perhaps surprisingly, our first main result shows that they are equivalent.

Theorem 2.9 *Let $\mathcal{F}$ satisfy Assumption 2.8. A numerical integrator Φ preserves $\mathcal{F}$-invariants if and only if it is $\mathcal{F}$-functionally equivariant.*

Proof ($\Rightarrow$) Suppose Φ preserves $\mathcal{F}$-invariants. Given $F \in \mathcal{F}(Y, Z)$, it follows from Assumption 2.8 that $G(y, z) = F(y) - z$ is in $\mathcal{F}(Y \times Z, Z)$. This is an invariant of the augmented vector field $g(y, z) = \big(f(y), F'(y) f(y)\big)$, since $G'(y, z) g(y, z) = F'(y) f(y) - F'(y) f(y) = 0$. Hence, preservation of $\mathcal{F}$-invariants implies $\Phi_g: (y_0, z_0) \mapsto (y_1, z_1)$ satisfies $G(y_1, z_1) = G(y_0, z_0)$, i.e., $F(y_1) - z_1 = F(y_0) - z_0$. In particular, $z_0 = F(y_0)$ implies $z_1 = F(y_1)$.

($\Leftarrow$) Conversely, suppose Φ is $\mathcal{F}$-functionally equivariant. If $F \in \mathcal{F}(Y, Z)$ is an invariant of $f \in \mathfrak{X}(Y)$, then the augmented vector field is $g = (f, 0)$, and $\mathcal{F}$-functional

equivariance implies $\Phi_g \colon \big(y_0, F(y_0) \big) \mapsto \big(y_1, F(y_1) \big)$. However, any constant functional $y \mapsto c \in Z$ is also in $\mathcal{F}(Y, Z)$ and has the same augmented vector field g, so $\Phi_g \colon (y_0, c) \mapsto (y_1, c)$ for all $c \in Z$. Hence, $F(y_1) = F(y_0)$. □

Corollary 2.10 *For B-series methods, the following statements hold:*

(a) *Every B-series method is affine functionally equivariant.*
(b) *B-series methods preserving quadratic invariants (e.g., Gauss–Legendre collocation methods) are quadratic functionally equivariant.*
(c) *No B-series method is cubic functionally equivariant.*

Proof This follows since B-series methods are affine equivariant and none preserves arbitrary cubic invariants (Chartier and Murua [9], Iserles, Quispel, and Tse [16]). □

2.2 Strong Equivariance vs. Functional Equivariance

There is a stronger notion of F-equivariance, based on a straightforward generalization of Definition 2.2 to nonlinear maps $F \colon Y \to U$. Two vector fields $f \in \mathfrak{X}(Y)$ and $g \in \mathfrak{X}(U)$ are F-related if $F'(y) f(y) = (g \circ F)(y)$ for all $y \in Y$, and Φ is F-equivariant if $F \circ \Phi_f = \Phi_g \circ F$ whenever this is the case.

To illustrate the distinction with functional equivariance, we now show that the implicit midpoint method is *not* quadratic equivariant in this stronger sense, although Corollary 2.10(b) tells us that it is quadratic functionally equivariant. Let $F \colon \mathbb{R} \to \mathbb{R}$, $y \mapsto y^2$, and observe that the vector fields

$$f(y) = -y, \qquad g(u) = -2u,$$

are F-related. Applying the implicit midpoint method with time step size $\Delta t = 1$ gives

$$y_1 = \frac{1}{3} y_0, \qquad u_1 = 0.$$

Since $u_1 \neq (y_1)^2$ for $y_0 \neq 0$, the method is not F-equivariant. On the other hand, applying the method to the augmented equation $\dot{z} = F'(y) f(y) = -2y^2$ with $z_0 = (y_0)^2$ gives

$$z_1 = (y_0)^2 - 2 \left(\frac{y_0 + y_1}{2} \right)^2 = (y_0)^2 - 2 \left(\frac{y_0 + \frac{1}{3} y_0}{2} \right)^2 = \frac{1}{9} (y_0)^2 = (y_1)^2,$$

which illustrates that the method is F-functionally equivariant.

Essentially, functional equivariance requires only that Φ commute with *particular* pairs of related vector fields, while strong equivariance requires that it commute with *all* such pairs.

2.3 Affine Equivariance and Closure under Differentiation

In addition to invariants and observables that depend on y itself, we are often interested in those that depend on variations of y. We say that η is a variation of y if $(y, \eta) \in Y \times Y$

satisfy

$$\dot{y} = f(y), \qquad \dot{\eta} = f'(y)\eta, \tag{2}$$

whose flow is the derivative of the flow of $f \in \mathfrak{X}(Y)$. An especially important example is the canonical symplectic form for Hamiltonian ODEs, which can be understood as a quadratic invariant depending on two variations of y.

Definition 2.11 A numerical method Φ is said to be *closed under differentiation* if the method applied to (2) is the derivative of Φ_f, i.e., $(y_1, \eta_1) = \big(\Phi_f(y_0), \Phi'_f(y_0)\eta_0\big)$.

Bochev and Scovel [5] showed that Runge–Kutta methods are closed under differentiation, from which it follows that those preserving quadratic invariants are symplectic integrators. The same argument can be applied to B-series methods, where closure under differentiation can be established by showing that it holds for all trees [15, Theorem VI.7.1]. Here, we present a new, tree-free proof that uses only affine equivariance, and which will readily generalize to additive and partitioned methods in Sect. 5.

Theorem 2.12 *Affine equivariant numerical integrators are closed under differentiation.*

Proof Given $f \in \mathfrak{X}(Y)$, consider the system

$$\dot{x} = f(x), \qquad \dot{y} = f(y),$$

corresponding to $f \times f \in \mathfrak{X}(Y \times Y)$. Since $f \times f$ is A-related to f, where A is either of the projections $(x, y) \mapsto x$ or $(x, y) \mapsto y$, it follows that $\Phi_{f \times f} = \Phi_f \times \Phi_f$. Now, let $\epsilon > 0$ and take $z = F(x, y) = (x - y)/\epsilon$, giving the augmented system

$$\dot{x} = f(x), \qquad \dot{y} = f(y), \qquad \dot{z} = \frac{f(x) - f(y)}{\epsilon}.$$

By Proposition 2.6, applying Φ to this system yields

$$x_1 = \Phi_f(x_0), \qquad y_1 = \Phi_f(y_0), \qquad z_1 = \frac{\Phi_f(x_0) - \Phi_f(y_0)}{\epsilon}.$$

Finally, let $x_0 = y_0 + \epsilon\eta_0$ and take the limit as $\epsilon \to 0$. $\qquad\qquad\square$

Corollary 2.13 *Let Φ be an affine equivariant numerical integrator preserving $\mathcal{F}$-invariants. Given $F \in \mathcal{F}(Y \times Y, Z)$, define $g \in \mathfrak{X}(Y \times Y \times Z)$ by*

$$g(y, \eta, z) = \Big(f(y), f'(y)\eta, F'(y, \eta)\big(f(y), f'(y)\eta\big)\Big).$$

Then $\Phi_g\big(y_0, \eta_0, F(y_0, \eta_0)\big) = \big(y_1, \eta_1, F(y_1, \eta_1)\big)$, where $y_1 = \Phi_f(y_0)$ and $\eta_1 = \Phi'_f(y_0)\eta_0$.

Proof Apply Theorems 2.9 and 2.12. $\qquad\qquad\square$

Remark 2.14 It is trivial to extend Corollary 2.13 to the case where F depends on two or more variations of y, e.g., $F = F(y, \xi, \eta)$ where $\dot{\xi} = f'(y)\xi$ and $\dot{\eta} = f'(y)\eta$.

2.4 Examples

Before discussing applications to conservation laws for PDEs, which will be the subject of Sect. 3, we first illustrate some examples of functional equivariance for numerical ODEs.

2.4.1 Hamiltonian Systems

Suppose Y is equipped with a Poisson bracket $\{\cdot, \cdot\}$. Given $H \colon Y \to \mathbb{R}$, the corresponding Hamiltonian vector field $f \in \mathfrak{X}(Y)$ is determined by the condition $\dot{F} = \{F, H\}$ for $F \colon Y \to \mathbb{R}$. That is, the augmented system (1) can be written

$$\dot{y} = f(y), \qquad \dot{z} = \{F, H\}(y).$$

Hence, if Φ is $\mathcal{F}$-functionally equivariant, then applying Φ to this system gives a "discrete version" of $\dot{F} = \{F, H\}$ for $F \in \mathcal{F}(Y, \mathbb{R})$. For a Runge–Kutta method, this has the form

$$F(y_1) = F(y_0) + \Delta t \sum_{i=1}^{s} b_i \{F, H\}(Y_i).$$

This holds for *any* Poisson bracket, not just the canonical bracket or those for which the Poisson tensor is constant. Preservation of $\mathcal{F}$-invariants is the special case where $\{F, H\} = 0$.

2.4.2 Canonical Hamiltonian Systems with Damping/Forcing

Let $Y = \mathbb{R}^{2n}$ with canonical coordinates $y = (q, p)$. Consider the system

$$\dot{q} = \nabla_p H(q, p), \qquad \dot{p} = -\nabla_q H(q, p) - cp, \qquad (3)$$

where H is a Hamiltonian and $c \geq 0$ a constant parameter. If H has the special form $H(q, p) = \frac{1}{2} p^T M^{-1} p + V(q)$, where M is a positive definite mass matrix and V is a potential energy function, then energy is dissipated according to $\frac{\mathrm{d}}{\mathrm{d}t} H(q, p) = -cp^T M^{-1} p$, and the parameter c dictates the rate of dissipation.

If V is also quadratic, then so is H, and hence any quadratic functionally equivariant method Φ yields a discrete version of this dissipation law. If Φ is a Runge–Kutta method preserving quadratic invariants, then this has the form

$$H(q_1, p_1) = H(q_0, p_0) - \Delta t \sum_{i=1}^{s} b_i c P_i^T M^{-1} P_i.$$

There is no reason to restrict to linear damping: if we replace the damping term $-cp$ in (3) by an arbitrary forcing term $\phi(q, p)$, then we obtain

$$H(q_1, p_1) = H(q_0, p_0) + \Delta t \sum_{i=1}^{s} b_i P_i^T M^{-1} \phi(Q_i, P_i), \qquad (4)$$

where the sum on the right-hand side approximates the work done by ϕ. When V is not quadratic, the identities above generally do not hold. However, since the kinetic energy functional $\frac{1}{2} p^T M^{-1} p$ is quadratic, we still get the weaker identity

$$\frac{1}{2} p_1^T M^{-1} p_1 = \frac{1}{2} p_0^T M^{-1} p_0 + \Delta t \sum_{i=1}^{s} b_i P_i^T M^{-1} \big[-\nabla V(Q_i) + \phi(Q_i, P_i) \big],$$

where the sum approximates work done by both conservative and non-conservative forces.

2.4.3 Monotone Observables

Suppose $F \in \mathcal{F}(Y, \mathbb{R})$ is such that $F'(y)f(y) \leq 0$, so $F(y)$ is monotone decreasing. If Φ is an $\mathcal{F}$-functionally equivariant Runge–Kutta method with $b_i \geq 0$, then

$$F(y_1) = F(y_0) + \Delta t \sum_{i=1}^{s} b_i F'(Y_i) f(Y_i) \leq F(y_0),$$

so F is also monotone decreasing along the numerical solution. Conversely, any method with this monotonicity property also preserves $\mathcal{F}$-invariants, and is thus $\mathcal{F}$-functionally equivariant, since F is an invariant when $\pm F$ are both monotone decreasing.

Remark 2.15 For Runge–Kutta methods, the additional condition $b_i \geq 0$ is needed to get monotonicity. Functional equivariance alone is not sufficient. We are not aware of a more general version of this condition for arbitrary B-series methods.

An immediate consequence is the known B-stability of Runge–Kutta methods preserving quadratic invariants with $b_i \geq 0$. If Y is a Hilbert space with inner product $\langle \cdot, \cdot \rangle$, consider
$$\dot{x} = f(x), \qquad \dot{y} = f(y),$$
on $Y \times Y$, and let $F(x, y) = \frac{1}{2} \|x - y\|^2$. Then $F'(x, y)\big(f(x), f(y)\big) = \langle x - y, f(x) - f(y) \rangle \leq 0$ implies $F(x_1, y_1) \leq F(x_0, y_0)$, i.e., $\|x_1 - y_1\| \leq \|x_0 - y_0\|$. This is precisely the condition for B-stability, cf. Butcher [8] , Burrage and Butcher [7].

Another immediate application is to the dissipative systems in Sect. 2.4.2, when H is quadratic. If ϕ is a dissipative force, in the sense that $p^T M^{-1} \phi(q, p) \leq 0$ for all $(q, p) \in \mathbb{R}^{2n}$, then (4) implies $H(q_1, p_1) \leq H(q_0, p_0)$, i.e., the quadratic energy is monotone decreasing along the numerical solution.

2.4.4 Symplectic and Conformal Symplectic Systems

Suppose that ω is a continuous bilinear form on Y. Let ξ and η each be variations of y, so that $(y, \xi, \eta) \in Y \times Y \times Y$ satisfy

$$\dot{y} = f(y), \qquad \dot{\xi} = f'(y)\xi, \qquad \dot{\eta} = f'(y)\eta.$$

Then $\omega(\xi, \eta)$ is a quadratic functional of this augmented system, evolving according to

$$\frac{\mathrm{d}}{\mathrm{d}t}\omega(\xi, \eta) = \omega\big(f'(y)\xi, \eta\big) + \omega\big(\xi, f'(y)\eta\big) = (L_f \omega)_y(\xi, \eta),$$

where $(L_f \omega)_y$ is the Lie derivative along f of ω at y [1, Theorem 6.4.1].

If Φ preserves quadratic invariants, then we may apply quadratic functional equivariance to describe the numerical evolution of ω. Taking $g \in \mathfrak{X}(Y \times Y \times Y \times \mathbb{R})$ to be

$$g(y, \xi, \eta, z) = \big(f(y), f'(y)\xi, f'(y)\eta, (L_f \omega)_y(\xi, \eta)\big),$$

it follows from Corollary 2.13 and Remark 2.14 that

$$\Phi_g\big(y_0, \xi_0, \eta_0, \omega(\xi_0, \eta_0)\big) = \big(y_1, \xi_1, \eta_1, \omega(\xi_1, \eta_1)\big), \tag{5}$$

where $y_1 = \Phi_f(y_0)$, $\xi_1 = \Phi'_f(y_0)\xi_0$, and $\eta_1 = \Phi'_f(y_0)\eta_0$. Furthermore, this implies

$$\omega(\xi_1, \eta_1) = \omega\big(\Phi'_f(y_0)\xi_0, \Phi'_f(y_0)\eta_0\big) = (\Phi_f^* \omega)_{y_0}(\xi_0, \eta_0),$$

where $(\Phi_f^* \omega)_{y_0}$ is the pullback of ω by Φ_f at y_0. For a Runge–Kutta method preserving quadratic invariants, (5) takes the form

$$\omega(\xi_1, \eta_1) = \omega(\xi_0, \eta_0) + \Delta t \sum_{i=1}^{s} b_i (L_f \omega)_{Y_i}(\Xi_i, \mathrm{H}_i), \tag{6}$$

and the sum on the right-hand side expresses the difference between $\Phi_{\Delta t f}^* \omega$ and ω at y_0.

In particular, suppose that ω is antisymmetric and nondegenerate, so that (Y, ω) is a symplectic vector space. If f is a symplectic vector field, satisfying $L_f \omega = 0$, then we recover the result of Bochev and Scovel [5] that if Φ preserves quadratic invariants, then $\Phi_f^* \omega = \omega$, i.e., Φ is a symplectic integrator. An interesting generalization is the case of *conformal symplectic* vector fields, satisfying $L_f \omega = -c\omega$ for some constant c, of which (3) is a canonical example; see McLachlan and Perlmutter [17]. In this case, (6) becomes

$$\omega(\xi_1, \eta_1) = \omega(\xi_0, \eta_0) - \Delta t \sum_{i=1}^{s} b_i c \omega(\Xi_i, \mathrm{H}_i),$$

which can be seen as an approximate conformal symplecticity relation. However, $\Phi_{\Delta t f}^* \omega$ generally does not equal $e^{-c\Delta t}\omega$ exactly unless $c = 0$; see McLachlan and Quispel [19, Example 7] for a counterexample when Φ is the implicit midpoint method.

Remark 2.16 The arguments above apply without modification if ω is a vector-valued bilinear form, i.e., a continuous bilinear map $Y \times Y \to Z$ for some Banach space Z.

3 Application to Conservation Laws in PDEs

In this section, we apply the general results of Sect. 2 to local conservation laws in time-evolution PDEs. We also consider discrete conservation laws in numerical PDEs, when semidiscretization in space is combined with numerical integration in time.

3.1 General Approach and Examples

Let $\dot{y} = f(y)$ correspond to a time-dependent system of PDEs on a domain Ω, where the Banach space Y is a function space (or product of function spaces) on Ω. Suppose that solutions satisfy a local conservation law, in the form

$$\dot{\rho} = -\operatorname{div} J, \tag{7}$$

where ρ and J depend on y. The notation is deliberately suggestive of Maxwell's equations, where ρ is charge density, J is current density, and (7) is local conservation of charge.

From Theorem 2.9 and Corollary 2.10, we immediately obtain a powerful general statement about preservation of local conservation laws under numerical integration. If $\rho = F(y)$, where $F \in \mathcal{F}(Y, Z)$ and Z is an appropriate space of densities, then $F'(y)f(y) = -\operatorname{div} J(y)$, and thus an $\mathcal{F}$-functionally equivariant integrator satisfies a discrete-time version of (7). For instance, a Runge–Kutta method preserving $\mathcal{F}$-invariants satisfies

$$\rho_1 = \rho_0 - \Delta t \sum_{i=1}^{s} b_i \operatorname{div} J(Y_i).$$

We note that, while ρ is required to be related to y by a functional in $\mathcal{F}$ (e.g., ρ is affine or quadratic in y), no such restriction is placed on J. In particular, all B-series methods preserve affine local conservation laws, while those preserving quadratic invariants also preserve quadratic local conservation laws. In the case of symplectic Runge–Kutta methods, Frasca-Caccia and Hydon [12, Section 3.1] recently proved this by a direct computation, whereas here it is seen as a particular instance of quadratic functional equivariance.

In addition to the differential form of the conservation law (7), we may also integrate over a compact subdomain $K \subset \Omega$ and apply the divergence theorem to get

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_K \rho = -\int_{\partial K} J \cdot \widehat{n}, \tag{8}$$

where $\widehat{n}$ denotes the outer unit normal to K. This may be seen as an integral form of the conservation law (7). In this case, if $\int_K \rho = F(y)$ with $F \in \mathcal{F}(Y, \mathbb{R})$, then an $\mathcal{F}$-functionally equivariant method satisfies a discrete-time version of (8). In the case of a Runge–Kutta method preserving $\mathcal{F}$-invariants, this has the form

$$\int_K \rho_1 = \int_K \rho_0 - \Delta t \sum_{i=1}^{s} b_i \int_{\partial K} J(Y_i) \cdot \widehat{n}.$$

Example 3.1 Maxwell's equations in $\mathbb{R}^3$ consist of the vector evolution equations

$$\dot{B} = -\operatorname{curl} E, \qquad \dot{D} = \operatorname{curl} H - J,$$

along with the scalar constraint equations

$$\operatorname{div} B = 0, \qquad \operatorname{div} D = \rho,$$

and the constitutive relations $D = \epsilon E$ and $B = \mu H$. Here, E and H are the electric and magnetic fields, D and B are the electric and magnetic flux densities, ϵ and μ are the electric permittivity and magnetic permeability tensors, and ρ and J are charge and current density.

Taking the divergence of the first evolution equation, we see that $\operatorname{div} B$ is a local invariant, so the constraint $\operatorname{div} B = 0$ is preserved by the evolution. Next, interpreting the constraint $\operatorname{div} D = \rho$ to define ρ as a function of D, we see that taking the divergence of the second evolution equation gives the local conservation law $\dot{\rho} = -\operatorname{div} J$. Since $\operatorname{div} B$ and $\operatorname{div} D$ are both linear in $y = (B, D)$, any B-series method will preserve the constraint $\operatorname{div} B = 0$, together with a discrete version of the conservation law relating ρ and J.

Example 3.2 Consider the nonlinear Schrödinger (NLS) equation,

$$i\dot{u} + \Delta u = \phi\big(|u|^2\big)u.$$

A direct computation shows that solutions satisfy

$$\frac{\partial}{\partial t}\frac{1}{2}|u|^2 = \Im(\bar{u}\, i\dot{u}) = \Im(-\bar{u}\Delta u) = \Im\big(-\operatorname{div}(\bar{u}\operatorname{grad} u)\big) = -\operatorname{div}\Im(\bar{u}\operatorname{grad} u),$$

which is a local conservation law for the quadratic functional $F(u) = \frac{1}{2}|u|^2$. Since the implicit midpoint method is quadratic functionally equivariant, it follows that this conservation law is preserved, in the sense that

$$\frac{1}{2}|u_1|^2 = \frac{1}{2}|u_0|^2 - \Delta t \operatorname{div} \Im\left(\frac{\bar{u}_0 + \bar{u}_1}{2}\operatorname{grad}\frac{u_0 + u_1}{2}\right).$$

More generally, for any quadratic functionally equivariant Runge–Kutta method,

$$\frac{1}{2}|u_1|^2 = \frac{1}{2}|u_0|^2 - \Delta t \sum_{i=1}^{s} b_i \operatorname{div} \Im(\bar{U}_i \operatorname{grad} U_i).$$

Example 3.3 Consider the wave equation $\ddot{u} = \Delta u$, written as the first-order system

$$\dot{u} = p, \qquad \dot{p} = \Delta u.$$

If $y = (u, p)$ is a solution, then

$$\frac{\partial}{\partial t} \frac{1}{2}(p^2 + |\text{grad } u|^2) = p\dot{p} + \text{grad } u \cdot \text{grad } \dot{u} = p\Delta u + \text{grad } u \cdot \text{grad } p = \text{div}(p \text{ grad } u),$$

which is a local conservation law for the quadratic functional $F(u, p) = \frac{1}{2}(p^2 + |\text{grad } u|^2)$. Hence, applying any quadratic functionally equivariant Runge–Kutta method gives

$$\frac{1}{2}((p_1)^2 + |\text{grad } u_1|^2) = \frac{1}{2}((p_0)^2 + |\text{grad } u_0|^2) + \Delta t \sum_{i=1}^{s} b_i \text{ div}(P_i \text{ grad } U_i).$$

3.2 Discrete Conservation Laws in Numerical PDEs

In practice, of course, we will not be applying numerical integrators to infinite-dimensional function spaces. Rather, we typically first semidiscretize in space (e.g., using a finite difference, finite volume, or finite-element scheme), yielding a system $\dot{y}_h = f_h(y_h)$ on a finite-dimensional vector space Y_h to which we can apply a numerical integrator.

Suppose the spatial discretization scheme is such that it preserves a semidiscrete conservation law $\dot{\rho}_h = -\text{div}_h J_h$, where div_h is some "discrete divergence" operator. Then it follows from the previous section that, if $\rho_h = F_h(y_h)$ for some $F_h = \mathcal{F}(Y_h, Z_h)$, then applying an $\mathcal{F}$-functionally equivariant numerical integrator yields a fully discrete conservation law corresponding to (7). We illustrate this with a few examples, which are semidiscretized versions of those considered in the previous section.

Example 3.4 Nédélec [24] introduced a finite-element semidiscretization of Maxwell's equations, in which E and B are approximated by piecewise-polynomial vector fields $E_h \in H(\text{curl}; \Omega)$ and $B_h \in H(\text{div}; \Omega)$. This method may be written as

$$\dot{B}_h = -\text{curl } E_h, \qquad \int_\Omega \dot{D}_h \cdot v_h = \int_\Omega (H_h \cdot \text{curl } v_h - J_h \cdot v_h),$$

where $D_h = \epsilon E_h$, $H_h = \mu^{-1} B_h$, and v_h is any vector field from the same space as E_h.

Taking the divergence of the first equation gives $\text{div } \dot{B}_h = 0$, so the constraint $\text{div } B_h = 0$ is preserved by the evolution. For the second, when $v_h = -\text{grad } \phi_h$ for a piecewise-polynomial scalar field ϕ_h, we get

$$-\int_\Omega \dot{D}_h \cdot \text{grad } \phi_h = \int_\Omega J_h \cdot \text{grad } \phi_h,$$

which we may write as $\text{div}_h \dot{D}_h = -\text{div}_h J_h$. Thus, taking $\rho_h = \text{div}_h D_h$ implies the semidiscrete charge conservation law $\dot{\rho}_h = -\text{div}_h J_h$. (See Berchenko-Kogan and

Stern [4] for a hybridization of Nédélec's method that preserves a stronger form of this conservation law, using div rather than div_h.) Since $\mathrm{div}\, B_h$ and $\mathrm{div}_h\, D_h$ are linear in $y_h = (B_h, D_h)$, any B-series method will preserve $\mathrm{div}\, B_h = 0$ exactly and give a discrete-time version of the charge conservation law relating ρ_h and J_h.

Example 3.5 For the one-dimensional NLS equation, the finite-difference semidiscretization

$$i\dot{u}_k + \frac{u_{k+1} - 2u_k + u_{k-1}}{h^2} = \phi\big(|u_k|^2\big)u_k$$

satisfies the semidiscrete local conservation law

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}|u_k|^2 = -\frac{1}{h}\left[\Im\left(\left(\frac{\bar{u}_k + \bar{u}_{k+1}}{2}\right)\left(\frac{u_{k+1} - u_k}{h}\right)\right) - \Im\left(\left(\frac{\bar{u}_{k-1} + \bar{u}_k}{2}\right)\left(\frac{u_k - u_{k-1}}{h}\right)\right)\right],$$

where the right-hand side is a difference of midpoint approximations to $\Im(\bar{u}\,\partial_x u)$. Hence, a discrete-time version of this conservation law is preserved by any B-series method that preserves quadratic invariants.

Example 3.6 For the one-dimensional wave equation, consider the finite-difference semidiscretization

$$\dot{u}_k = p_k, \qquad \dot{p}_k = \frac{u_{k+1} - 2u_k + u_{k-1}}{h^2}.$$

If we define

$$\rho_k = \frac{1}{2}p_k^2 + \frac{1}{4}\left(\frac{u_{k+1} - u_k}{h}\right)^2 + \frac{1}{4}\left(\frac{u_k - u_{k-1}}{h}\right)^2,$$

which is a finite-difference approximation to $\frac{1}{2}\big(p^2 + (\partial_x u)^2\big)$, then a short calculation gives the semidiscrete conservation law

$$\dot{\rho}_k = \frac{1}{h}\left[\left(\frac{p_k + p_{k+1}}{2}\right)\left(\frac{u_{k+1} - u_k}{h}\right) - \left(\frac{p_{k-1} + p_k}{2}\right)\left(\frac{u_k - u_{k-1}}{h}\right)\right],$$

where the right-hand side is a difference of midpoint approximations to $p\,\partial_x u$. As in the previous example, a discrete-time version of this conservation law is therefore preserved by any B-series method that preserves quadratic invariants.

3.3 Remarks on Quadratic Conservation Laws Arising from Point Symmetries

Conservation laws with quadratic densities are common in partial differential and differential-difference equations because of their association with linear symmetries of Hamiltonian PDEs. (See, e.g., Olver [25].) However, not all such symmetries are easily preserved under semidiscretization. We focus here on affine point symmetries, those arising from actions on the field variables.

For example, the one-dimensional NLS equation may be written in the form

$$i\dot{u} = \frac{\delta}{\delta\bar{u}}\int_\Omega\Big(|\partial_x u|^2 + V\big(|u|^2\big)\Big),$$

where $V' = \phi$ and $\delta/\delta\bar{u}$ is the variational derivative with respect to $\bar{u}$. The integrand $\mathcal{H} = |\partial_x u|^2 + V(|u|^2)$ is called the *Hamiltonian density*. Observe that $\mathcal{H}$ is invariant under the diagonal $U(1)$ action $(u, \partial_x u) \mapsto (e^{i\alpha}u, e^{i\alpha}\partial_x u)$, where $\alpha \in \mathbb{R} \cong \mathfrak{u}(1)$. This point symmetry leads to the local conservation law for $\rho = \frac{1}{2}|u|^2$ in Example 3.2. More generally, any Hamiltonian density of the form $\mathcal{H} = \mathcal{H}(|u|^2, |\partial_x u|^2, \bar{u}\,\partial_x u)$ has the same point symmetry, and hence $i\dot{u} = \frac{\delta}{\delta\bar{u}}\int_\Omega \mathcal{H}$ has a local conservation law for $\rho = \frac{1}{2}|u|^2$.

Similarly, the one-dimensional semidiscretized NLS equation in Example 3.6 can be written

$$i\dot{u}_k = \frac{\partial}{\partial\bar{u}_k}\sum_\ell\left(\left|\frac{u_{\ell+1} - u_\ell}{h}\right|^2 + \frac{V(|u_\ell|^2) + V(|u_{\ell+1}|^2)}{2}\right),$$

where the summand can be viewed as a discrete Hamiltonian density $\mathcal{H}_h$. The invariance of $\mathcal{H}_h$ under the point symmetry $u_\ell \mapsto e^{i\alpha}u_\ell$, $u_{\ell+1} \mapsto e^{i\alpha}u_{\ell+1}$, yields the semidiscrete local conservation law for $\rho_k = \frac{1}{2}|u_k|^2$ obtained in Example 3.6. More generally, we get such a local conservation law whenever the discrete Hamiltonian density has the form $\mathcal{H}_h = \mathcal{H}_h(|u_\ell|^2, |u_{\ell+1}|^2, \bar{u}_\ell u_{\ell+1})$.

A related example involves orthogonal (rather than unitary) point symmetry. Suppose $u(t, x)$ and its conjugate momentum $p(t, x)$ both take values in $\mathbb{R}^3$, and let $A \in O(3)$ act by $z = (u, p, \partial_x u, \partial_x p) \mapsto (Au, Ap, A\partial_x u, A\partial_x p)$. Then any Hamiltonian density that depends only on the 10 invariants $z_i \cdot z_j$, $1 \leq i \leq j \leq 4$, is $O(3)$ invariant and thus has a local conservation law for $\rho = u \times p$. Like the $U(1)$ point symmetry discussed above, this $O(3)$ point symmetry is preserved under a wide class of lattice semidiscretizations, which have corresponding semidiscrete quadratic conservation laws.

By contrast with point symmetries, symmetries that involve spatial translations are typically broken by semidiscretization. However, special semidiscretizations can be constructed that preserve versions of the associated conservation laws, although these are generally not symplectic. An example is provided by the Korteweg–de Vries equation

$$\partial_t u = \partial_x(\alpha u^2) + \nu\partial_x^3 u,$$

which has a local conservation law with $\rho = u^2$. The semidiscretization

$$\dot{u}_k = \frac{\alpha}{2h}\left[\theta(u_{k+1}^2 - u_{k-1}^2) + 2(1-\theta)u_k(u_{k+1} - u_{k-1})\right] + \frac{\nu}{2h^3}(u_{k+2} - 2u_{k+1} + 2u_{k-1} - u_{k-2})$$

has a semidiscrete conservation law with density $\rho_k = u_k^2$ only for the parameter $\theta = 2/3$ (Ascher and McLachlan [3]).

Frasca-Caccia and Hydon [12] give general techniques for constructing finite-difference semidiscretizations preserving several local conservation laws—linear, quadratic, or otherwise—with many examples. When such methods are used in conjunction with B-series methods for time integration, it follows from Theorem 2.9 and Corollary 2.10 that affine local conservation laws are always preserved in a discrete

sense, while quadratic local conservation laws are preserved by any B-series method that preserves quadratic invariants.

4 Multisymplectic Integrators

In this section, we apply the foregoing theory to the multisymplectic conservation law for canonical Hamiltonian PDEs and its preservation by numerical integrators. Since this is a quadratic local conservation law depending on variations of solutions, it follows that B-series methods preserving quadratic invariants also preserve a discrete-time version of the multisymplectic conservation law. Furthermore, we discuss techniques for spatial semidiscretization that preserve a semidiscrete multisymplectic conservation law, reviewing some known results for finite-difference semidiscretization and introducing new results for finite-element semidiscretization. Consequently, when such methods are used in conjunction with B-series methods preserving quadratic invariants, the resulting method will satisfy a fully discrete multisymplectic conservation law.

4.1 Canonical Hamiltonian PDEs

Before discussing the canonical Hamiltonian formalism for time-evolution PDEs, we first quickly recall the stationary (time-independent) case, following the treatment in McLachlan and Stern [22].

Given a spatial domain $\Omega \subset \mathbb{R}^m$ with coordinates $x = (x^1, \ldots, x^m)$, let $u \colon \Omega \to \mathbb{R}^n$ and $\sigma \colon \Omega \to \mathbb{R}^{mn}$ be unknown fields. The *de Donder–Weyl equations* [11, 30] for a Hamiltonian $H \colon \Omega \times \mathbb{R}^n \times \mathbb{R}^{mn} \to \mathbb{R}$, $H = H(x, u, \sigma)$, are

$$\partial_\mu u^i = \frac{\partial H}{\partial \sigma_i^\mu}, \qquad -\partial_\mu \sigma_i^\mu = \frac{\partial H}{\partial u^i}, \tag{9}$$

where $\mu = 1, \ldots, m$ and $i = 1, \ldots, n$. Here and henceforth, we use the Einstein index convention of summing over repeated indices; for instance, $\partial_\mu \sigma_i^\mu$ has an implied sum over μ and therefore corresponds to the divergence of the vector field σ_i.

Now, for time-dependent problems, we let $u = u^i(t, x)$ and $\sigma = \sigma_i^\mu(t, x)$ depend on $t \in (t_0, t_1)$, and we introduce an additional unknown field $p = p_i(t, x)$. The de Donder–Weyl equations for $H \colon (t_0, t_1) \times \Omega \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{mn}$, $H = H(t, x, u, p, \sigma)$, are then given by

$$\dot{u}^i = \frac{\partial H}{\partial p_i}, \qquad \partial_\mu u^i = \frac{\partial H}{\partial \sigma_i^\mu}, \qquad -(\dot{p}_i + \partial_\mu \sigma_i^\mu) = \frac{\partial H}{\partial u^i}. \tag{10}$$

Note that (10) is simply (9) in $(m+1)$-dimensional spacetime, where we have adopted the special notation $t = x^0$ and $p_i = \sigma_i^0$. Moreover, the special case $m = 0$ recovers canonical Hamiltonian mechanics on $\mathbb{R}^{2n}$.

For $m > 0$, the de Donder–Weyl equations are not in the form $\dot{y} = f(y)$, since we have expressions for $\dot{u}$ and $\dot{p}$ but not $\dot{\sigma}$. To deal with this, we assume that the second

equation of (10) defines σ as an implicit function of $t, x, u, p,$ and grad u. By the implicit function theorem, this is true (at least locally) if the $mn \times mn$ matrix $\partial^2 H / (\partial \sigma_i^\mu \partial \sigma_j^\nu)$ is nondegenerate. Therefore, we may eliminate the second equation and substitute this expression for σ into the other two equations. Assuming the Hamiltonian does not depend on t, this gives a system of the form $\dot{y} = f(y)$ with $y = (u, p)$.

Example 4.1 Let $n = 1$, so that u and p are scalar fields and σ is a vector field on Ω, and take $H = \frac{1}{2}(p^2 - |\sigma|^2)$. Then the de Donder–Weyl equations are

$$\dot{u} = p, \qquad \text{grad } u = -\sigma, \qquad -(\dot{p} + \text{div } \sigma) = 0.$$

Eliminating the second equation and substituting $\sigma = -\text{grad } u$ into the third, we obtain the first-order form of the wave equation with $y = (u, p)$, as in Example 3.3.

4.2 The Multisymplectic Conservation Law

For Hamiltonian ODEs, the symplectic conservation law is a statement about variations of solutions to Hamilton's equations. Similarly, for Hamiltonian PDEs, the multisymplectic conservation law is a statement about variations of solutions to the de Donder–Weyl equations.

Definition 4.2 Let (u, p, σ) be a solution to (10). A *(first) variation* of (u, p, σ) is a solution (v, r, τ) to the linearized problem

$$\dot{v}^i = \frac{\partial^2 H}{\partial p_i \partial u^j} v^j + \frac{\partial^2 H}{\partial p_i \partial p_j} r_j + \frac{\partial^2 H}{\partial p_i \partial \sigma_j^\nu} \tau_j^\nu,$$

$$\partial_\mu v^i = \frac{\partial^2 H}{\partial \sigma_i^\mu \partial u^j} v^j + \frac{\partial^2 H}{\partial \sigma_i^\mu \partial p_j} r_j + \frac{\partial^2 H}{\partial \sigma_i^\mu \partial \sigma_j^\nu} \tau_j^\nu,$$

$$-(\dot{r}_i + \partial_\mu \tau_i^\mu) = \frac{\partial^2 H}{\partial u^i \partial u^j} v^j + \frac{\partial^2 H}{\partial u^i \partial p_j} r_j + \frac{\partial^2 H}{\partial u^i \partial \sigma_j^\nu} \tau_j^\nu,$$

where the Hessians on the right-hand side are evaluated at (t, x, u, p, σ).

On the space $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{mn} \ni (u, p, \sigma)$, we now define the canonical 2-forms $\omega^0 = du^i \wedge dp_i$ and $\omega^\mu = du^i \wedge d\sigma_i^\mu$ for $\mu = 1, \ldots, m$. The *multisymplectic conservation law* states that, for any pair of variations (v, r, τ) and (v', r', τ'), we have

$$\partial_t \left(\omega^0 \big((v, r, \tau), (v', r', \tau') \big) \right) = -\partial_\mu \left(\omega^\mu \big((v, r, \tau), (v', r', \tau') \big) \right),$$

that is,

$$\partial_t (v^i r_i' - v'^i r_i) = -\partial_\mu (v^i \tau_i'^\mu - v'^i \tau_i^\mu).$$

The proof is simply a calculation, using the symmetry of the Hessian. We abbreviate the multisymplectic conservation law as

$$\dot{\omega}^0 = -\partial_\mu \omega^\mu, \tag{11}$$

with the understanding that both sides are evaluated on variations of solutions to (10). In the special case $m = 0$, we recover the usual symplectic conservation law for Hamiltonian ODEs. As with the conservation laws in Sect. 3.1, we may also integrate (11) over a compact subdomain $K \subset \Omega$ and apply the divergence theorem to get

$$\int_K \dot{\omega}^0 \, \mathrm{d}^m x = - \int_{\partial K} \omega^\mu \, \mathrm{d}^{m-1} x_\mu, \tag{12}$$

which is an integral form of the multisymplectic conservation law. Here, $\mathrm{d}^m x := \mathrm{d} x^1 \wedge \cdots \wedge \mathrm{d} x^m$ is the standard Euclidean volume form on $\mathbb{R}^m$ and $\mathrm{d}^{m-1} x_\mu := \iota_{e_\mu} \mathrm{d}^m x$ is its interior product with the μth standard basis vector. Again, we interpret (12) to mean that the equality holds when both sides are evaluated on arbitrary variations of solutions.

Remark 4.3 If Ω is compact, and boundary conditions are chosen so that $\omega^\mu \, \mathrm{d}^{m-1} x_\mu = 0$ on $\partial \Omega$, then taking $K = \Omega$ in (12) gives $\int_\Omega \dot{\omega}^0 \, \mathrm{d}^m x = 0$. This may be interpreted as invariance of the symplectic form $\omega = \int_\Omega \omega^0 \, \mathrm{d}^m x$ on the infinite-dimensional phase space Y.

4.3 Discrete-Time Multisymplectic Conservation Laws for Numerical Integrators

As before, assume that the Hamiltonian does not depend on t and that σ is an implicit function of the remaining variables, so that (10) can be reduced to $\dot{y} = f(y)$ with $y = (u, p)$. It follows that variations evolve according to $\dot{\eta} = f'(y)\eta$ with $\eta = (v, r)$, where $\tau = \sigma'(y)\eta$. Hence, (11) may be seen as a quadratic conservation law involving variations of y, and the results of Sect. 2 immediately apply to give the following.

Theorem 4.4 *Suppose that* (10) *can be written as* $\dot{y} = f(y)$ *with* $y = (u, p)$*, where* σ *is an implicit function of the other variables. Let* $\xi = (v, r)$ *and* $\eta = (v', r')$*, with* $\tau = \sigma'(y)\xi$ *and* $\tau' = \sigma'(y)\eta$*, and define the augmented vector field*

$$g(y, \xi, \eta, z) = \left(f(y), \, f'(y)\xi, \, f'(y)\eta, \, -\partial_\mu \left(\omega^\mu \left((v, r, \tau), (v', r', \tau') \right) \right) \right).$$

If Φ is an affine equivariant method preserving quadratic invariants, then

$$\Phi_g \left(y_0, \xi_0, \eta_0, \omega^0 \left((v_0, r_0, \tau_0), (v'_0, r'_0, \tau'_0) \right) \right) = \left(y_1, \xi_1, \eta_1, \omega^0 \left((v_1, r_1, \tau_1), (v'_1, r'_1, \tau'_1) \right) \right),$$

where $y_1 = \Phi_f(y_0)$, $\xi_1 = \Phi'_f(y_0)\xi_0$, and $\eta_1 = \Phi'_f(y_0)\eta_0$.

Proof The key observation is that $\omega^0 \left((v, r, \tau), (v', r', \tau') \right) = v^i r'_i - v'^i r_i$ is quadratic in ξ and η alone, so it is not affected by the (possibly nonlinear) dependence of σ and its variations on the other variables. Hence, the result follows from (5) and Remark 2.16. □

Corollary 4.5 *For a Runge–Kutta method preserving quadratic invariants, we have*

$$\omega^0\big((v_1, r_1, \tau_1), (v_1', r_1', \tau_1')\big)$$
$$= \omega^0\big((v_0, r_0, \tau_0), (v_0', r_0', \tau_0')\big) - \Delta t \sum_{i=1}^{s} b_i \partial_\mu \Big(\omega^\mu\big((V_i, R_i, T_i), (V_i', R_i', T_i')\big)\Big).$$

This may be written equivalently as

$$(\mathrm{d}u_1)^j \wedge (\mathrm{d}p_1)_j = (\mathrm{d}u_0)^j \wedge (\mathrm{d}p_0)_j - \Delta t \sum_{i=1}^{s} b_i \partial_\mu \big((\mathrm{d}U_i)^j \wedge (\mathrm{d}\Sigma_i)_j^\mu\big).$$

4.4 Multisymplectic Semidiscretization on Rectangular Grids

If Ω is a Cartesian product of intervals, equipped with a rectangular finite-difference grid, there is a substantial literature on spatial semidiscretization such that a semidiscrete multisymplectic conservation law holds. We refer the reader in particular to the following (non-exhaustive) list of references: Reich [26], Bridges and Reich [6], Ryland and McLachlan [27], McLachlan, Ryland, and Sun [21]. These semidiscretization schemes generally apply a symplectic Runge–Kutta or partitioned Runge–Kutta method in each of the spatial directions. In light of Sect. 2, the semidiscrete multisymplectic conservation law may be seen as resulting from m applications of (6) or its generalization to partitioned methods in Sect. 5.

In one dimension of space, Sun and Xing [29] have recently investigated multisymplectic semidiscretization using discontinuous Galerkin finite-element methods.

4.5 Multisymplectic Semidiscretization with Hybrid Finite Element Methods

In McLachlan and Stern [22], we developed a framework for multisymplectic discretization of time-independent Hamiltonian PDEs by hybrid finite element methods, including hybridizable discontinuous Galerkin methods (cf. Cockburn, Gopalakrishnan, and Lazarov [10]). In this section, we show that those same methods may be used for semidiscretization of time-dependent Hamiltonian PDEs, and that a semidiscrete multisymplectic conservation law holds. Consequently, when combined with a symplectic numerical integrator for time discretization, the resulting method satisfies a fully discrete multisymplectic conservation law in spacetime. Unlike the methods discussed in the previous section, these methods may be applied to unstructured meshes on non-rectangular domains.

Suppose that $\Omega \subset \mathbb{R}^m$ is polyhedral, and let $\mathcal{T}_h$ be a simplicial triangulation of Ω by m-simplices $K \in \mathcal{T}_h$, where $\mathcal{E}_h = \bigcup_{K \in \mathcal{T}_h} \partial K$ denotes the set of $(m-1)$-dimensional facets. We specify finite-element spaces

$$V(K) \subset \big[H^2(K)\big]^n, \qquad V := \prod_{K \in \mathcal{T}_h} V(K),$$

$$\Sigma(K) \subset \left[H^1(K)\right]^{mn}, \qquad \Sigma := \prod_{K \in \mathcal{T}_h} \Sigma(K),$$

along with spaces of approximate boundary traces on $\mathcal{E}_h$,

$$\widehat{V} \subset \left[L^2(\mathcal{E}_h)\right]^n, \qquad \widehat{V}_0 := \left\{\widehat{v} \in \widehat{V} : \widehat{v}|_{\partial\Omega} = 0\right\}.$$

The de Donder–Weyl equations (10) are then approximated by the weak problem: Find $\left(u(t), \sigma(t), p(t), \widehat{u}(t)\right) \in V \times \Sigma \times V^* \times \widehat{V}$ satisfying

$$\int_K \dot{u}^i r_i \, \mathrm{d}^m x = \int_K \frac{\partial H}{\partial p_i} r_i \, \mathrm{d}^m x, \quad \forall r \in V^*(K), \tag{13a}$$

$$0 = \int_K \left(u^i \partial_\mu \tau_i^\mu + \frac{\partial H}{\partial \sigma_i^\mu} \tau_i^\mu\right) \mathrm{d}^m x - \int_{\partial K} \widehat{u}^i \tau_i^\mu \, \mathrm{d}^{m-1} x_\mu, \quad \forall \tau \in \Sigma(K), \tag{13b}$$

$$\int_K \dot{p}_i v^i \, \mathrm{d}^m x = \int_K \left(\sigma_i^\mu \partial_\mu v^i - \frac{\partial H}{\partial u^i} v^i\right) \mathrm{d}^m x - \int_{\partial K} \widehat{\sigma}_i^\mu v^i \, \mathrm{d}^{m-1} x_\mu, \quad \forall v \in V(K), \tag{13c}$$

for all $K \in \mathcal{T}_h$, together with the *conservativity condition*

$$0 = \sum_{K \in \mathcal{T}_h} \int_{\partial K} \widehat{\sigma}_i^\mu \widehat{v}^i \, \mathrm{d}^{m-1} x_\mu, \quad \forall \widehat{v} \in \widehat{V}_0. \tag{13d}$$

Here, $\widehat{\sigma}$ is determined by $u, \sigma, \widehat{u}$ through a specified numerical flux function; see Cockburn, Gopalakrishnan, and Lazarov [10], McLachlan and Stern [22] for further details. The Eqs. (13a)–(13c) are derived by multiplying (10) by test functions, integrating by parts over K, and replacing the boundary traces of u and σ by the approximate traces $\widehat{u}$ and $\widehat{\sigma}$. Under appropriate nondegeneracy assumptions, the Eqs. (13a) and (13c) define the dynamics of $y_h = (u, p)$ on $Y_h := V \times V^*$, where $\sigma, \widehat{u}$, and $\widehat{\sigma}$ are implicit functions of y_h.

We may then consider variations of solutions to (13), along with a corresponding semidiscrete multisymplectic conservation law in the integral form (12). The following is a straightforward generalization of Lemma 2 in [22].

Theorem 4.6 *If* (13a)–(13c) *hold on* $K \in \mathcal{T}_h$, *then*

$$\int_K \partial_t(\mathrm{d}u^i \wedge \mathrm{d}p_i) \, \mathrm{d}^m x = -\int_{\partial K} (\mathrm{d}\widehat{u}^i \wedge \mathrm{d}\widehat{\sigma}_i^\mu) \, \mathrm{d}^{m-1} x_\mu$$
$$+ \int_{\partial K} \left(\mathrm{d}(\widehat{u}^i - u^i) \wedge \mathrm{d}(\widehat{\sigma}_i^\mu - \sigma_i^\mu)\right) \mathrm{d}^{m-1} x_\mu.$$

Consequently, the semidiscrete multisymplectic conservation law

$$\int_K \partial_t(\mathrm{d}u^i \wedge \mathrm{d}p_i) \, \mathrm{d}^m x = -\int_{\partial K} (\mathrm{d}\widehat{u}^i \wedge \mathrm{d}\widehat{\sigma}_i^\mu) \, \mathrm{d}^{m-1} x_\mu \tag{14}$$

holds on $K \in \mathcal{T}_h$ if and only if $\int_{\partial K} \left(\mathrm{d}(\widehat{u}^i - u^i) \wedge \mathrm{d}(\widehat{\sigma}_i^\mu - \sigma_i^\mu) \right) \mathrm{d}^{m-1} x_\mu = 0$.

Proof We may rewrite (13a)–(13c) as

$$\int_K \dot{u}^i \, \mathrm{d}p_i \, \mathrm{d}^m x = \int_K \frac{\partial H}{\partial p_i} \, \mathrm{d}p_i \, \mathrm{d}^m x,$$

$$0 = \int_K \left(u^i \, \mathrm{d}(\partial_\mu \sigma_i^\mu) + \frac{\partial H}{\partial \sigma_i^\mu} \, \mathrm{d}\sigma_i^\mu \right) \mathrm{d}^m x - \int_{\partial K} \widehat{u}^i \, \mathrm{d}\sigma_i^\mu \, \mathrm{d}^{m-1} x_\mu,$$

$$\int_K \dot{p}_i \, \mathrm{d}u^i \, \mathrm{d}^m x = \int_K \left(\sigma_i^\mu \mathrm{d}(\partial_\mu u^i) - \frac{\partial H}{\partial u^i} \, \mathrm{d}u^i \right) \mathrm{d}^m x - \int_{\partial K} \widehat{\sigma}_i^\mu \, \mathrm{d}u^i \, \mathrm{d}^{m-1} x_\mu.$$

Adding the first two equations, subtracting the third, and taking the exterior derivative on both sides, we get

$$\int_K \partial_t (\mathrm{d}u^i \wedge \mathrm{d}p_i) \, \mathrm{d}^m x$$

$$= \int_K \left(\partial_\mu (\mathrm{d}u^i \wedge \mathrm{d}\sigma_i^\mu) + \mathrm{dd}H \right) \mathrm{d}^m x - \int_{\partial K} (\mathrm{d}\widehat{u}^i \wedge \mathrm{d}\sigma_i^\mu + \mathrm{d}u^i \wedge \mathrm{d}\widehat{\sigma}_i^\mu) \, \mathrm{d}^{m-1} x_\mu$$

$$= \int_{\partial K} (\mathrm{d}u^i \wedge \mathrm{d}\sigma_i^\mu - \mathrm{d}\widehat{u}^i \wedge \mathrm{d}\sigma_i^\mu - \mathrm{d}u^i \wedge \mathrm{d}\widehat{\sigma}_i^\mu) \, \mathrm{d}^{m-1} x_\mu$$

$$= -\int_{\partial K} (\mathrm{d}\widehat{u}^i \wedge \mathrm{d}\widehat{\sigma}_i^\mu) \, \mathrm{d}^{m-1} x_\mu + \int_{\partial K} \left(\mathrm{d}(\widehat{u}^i - u^i) \wedge \mathrm{d}(\widehat{\sigma}_i^\mu - \sigma_i^\mu) \right) \mathrm{d}^{m-1} x_\mu,$$

where the second equality uses $\mathrm{dd}H = 0$ and the divergence theorem. $\qquad \square$

In Section 4 of [22], it is proved that several families of hybrid finite-element methods, including hybridized mixed methods (RT-H and BDM-H), nonconforming methods (NC-H), discontinuous Galerkin methods (LDG-H and IP-H), and continuous Galerkin methods (CG-H) satisfy the condition $\int_{\partial K} \left(\mathrm{d}(\widehat{u}^i - u^i) \wedge \mathrm{d}(\widehat{\sigma}_i^\mu - \sigma_i^\mu) \right) \mathrm{d}^{m-1} x_\mu = 0$ of Theorem 4.6. Therefore, when these methods are applied to (13), they satisfy the semidiscrete multisymplectic conservation law (14) on each $K \in \mathcal{T}_h$.

If the numerical flux satisfies the so-called *strong conservativity condition* $[\![\widehat{\sigma}]\!] = 0$, which is stronger than (13d), then the multisymplectic conservation law (14) may also be strengthened so that it holds for arbitrary unions of simplices. This holds for all of the methods mentioned in the previous paragraph *except* CG-H. The following is a straightforward generalization of Theorem 3 in [22].

Theorem 4.7 *If a strongly conservative method satisfies* (14), *then for all* $\mathcal{K} \subset \mathcal{T}_h$,

$$\int_{\bigcup \mathcal{K}} \partial_t (\mathrm{d}u^i \wedge \mathrm{d}p_i) \, \mathrm{d}^m x = -\int_{\partial(\overline{\bigcup K})} (\mathrm{d}\widehat{u}^i \wedge \mathrm{d}\widehat{\sigma}_i^\mu) \, \mathrm{d}^{m-1} x_\mu.$$

Proof Sum (14) over $K \in \mathcal{K}$, using $[\![\widehat{\sigma}]\!] = 0$ to cancel the contributions of internal facets. $\qquad \square$

Remark 4.8 In the situation considered in Remark 4.3, taking $\mathcal{K} = \mathcal{T}_h$ implies conservation of the symplectic form $\int_\Omega du^i \wedge dp_i \, d^m x$ on Y_h. This generalizes a result of Sánchez et al. [28], which states that semidiscretization of the acoustic wave equation by LDG-H is symplectic.

5 Generalization to Additive and Partitioned Methods

In the preceding sections, we have developed a theory of functional equivariance for a class of numerical integrators, including B-series methods, and applied it to local conservation laws for PDEs. This section extends the functional equivariance theory from Sect. 2 to two larger classes of numerical integrators: additive methods and partitioned methods. It follows that, when these methods are applied to PDEs satisfying local conservation laws, the results of Sects. 3 and 4 may also be extended to these classes of methods.

5.1 Additive Methods

We now consider integrators applied to a vector field $f \in \mathfrak{X}(Y)$ after it has been additively decomposed as $f = f^{[1]} + \cdots + f^{[N]}$. Specifically, we have in mind additive Runge–Kutta and NB-series methods (cf. Aráujo, Murua, and Sanz-Serna [2]), as well as splitting and composition methods (cf. McLachlan and Quispel [20]).

Denote the application of a method Φ to a decomposed vector field $f = f^{[1]} + \cdots + f^{[N]}$ by $\Phi_{f^{[1]},\ldots,f^{[N]}}$. By an additive numerical integrator, we mean the entire collection of maps $\Phi = \left\{ \Phi_{f^{[1]},\ldots,f^{[N]}} : f^{[1]},\ldots,f^{[N]} \in \mathfrak{X}(Y), \; Y \text{ a Banach space} \right\}$. We begin by extending the definitions of affine equivariance and functional equivariance to such methods.

Definition 5.1 An additive numerical integrator Φ is *N-affine equivariant* if $A \circ \Phi_{f^{[1]},\ldots,f^{[N]}} = \Phi_{g^{[1]},\ldots,g^{[N]}} \circ A$ whenever $f^{[\nu]} \in \mathfrak{X}(Y)$ and $g^{[\nu]} \in \mathfrak{X}(U)$ are A-related for all $\nu = 1,\ldots,N$, all affine maps $A\colon Y \to U$, and all Banach spaces Y and U.

Definition 5.2 Given a Gâteaux differentiable map $F\colon Y \to Z$ and $f^{[1]},\ldots,f^{[N]} \in \mathfrak{X}(Y)$, define $g^{[1]},\ldots,g^{[N]} \in \mathfrak{X}(Y \times Z)$ by $g^{[\nu]}(y,z) = \left(f^{[\nu]}(y), F'(y) f^{[\nu]}(y) \right)$ for $\nu = 1,\ldots,N$. We say that an additive numerical integrator Φ is *F-functionally equivariant* if $(\mathrm{id}, F) \circ \Phi_{f^{[1]},\ldots,f^{[N]}} = \Phi_{g^{[1]},\ldots,g^{[N]}} \circ (\mathrm{id}, F)$ for all $f^{[1]},\ldots,f^{[N]} \in \mathfrak{X}(Y)$ and *$\mathcal{F}$-functionally equivariant* if this holds for all $F \in \mathcal{F}(Y, Z)$ and all Banach spaces Y and Z.

Proposition 5.3 *Every N-affine equivariant method is affine functionally equivariant.*

Proof The proof is essentially identical to that for Proposition 2.6. If F is affine, then so is (id, F), and the vector fields $f^{[\nu]}$ and $g^{[\nu]}$ are (id, F)-related for all $\nu = 1,\ldots,N$. $\square$

Example 5.4 (additive Runge–Kutta methods)   An s-stage additive Runge–Kutta (ARK) method has the form

$$Y_i = y_0 + \Delta t \sum_{\nu=1}^{N} \sum_{j=1}^{s} a_{ij}^{[\nu]} f^{[\nu]}(Y_j), \qquad i = 1, \ldots, s,$$

$$y_1 = y_0 + \Delta t \sum_{\nu=1}^{N} \sum_{i=1}^{s} b_i^{[\nu]} f^{[\nu]}(Y_i),$$

and F-functional equivariance is the condition

$$F(y_1) = F(y_0) + \Delta t \sum_{\nu=1}^{N} \sum_{i=1}^{s} b_i^{[\nu]} F'(Y_i) f^{[\nu]}(Y_i).$$

If F is an invariant, then we have $F'(Y_i) f(Y_i) = 0$ but generally $F'(Y_i) f^{[\nu]}(Y_i) \neq 0$ for $N > 1$, so the sum on the right-hand side need not vanish. However, if $b_i^{[\nu]} = b_i$ is independent of ν, then it does vanish, and we obtain $F(y_1) = F(y_0)$ as in Example 2.5. This illustrates that an ARK method may be functionally equivariant but not invariant preserving (even for affine maps) unless some additional condition is satisfied.

Proposition 5.5 *Additive Runge–Kutta methods are N-affine equivariant. Furthermore, an ARK method preserves affine invariants if $b_i^{[\nu]} = b_i$ is independent of ν.*

Proof Suppose $f^{[\nu]}$ and $g^{[\nu]}$ are A-related for $\nu = 1, \ldots, N$. Then

$$A(Y_i) = A(y_0) + A'(Y_i - y_0)$$

$$= A(y_0) + \Delta t \sum_{\nu=1}^{N} \sum_{j=1}^{s} a_{ij}^{[\nu]} (A' \circ f^{[\nu]})(Y_j)$$

$$= A(y_0) + \Delta t \sum_{\nu=1}^{N} \sum_{j=1}^{s} a_{ij}^{[\nu]} g^{[\nu]}(A(Y_j)),$$

for $i = 1, \ldots, s$, and similarly,

$$A(y_1) = A(y_0) + \Delta t \sum_{\nu=1}^{N} \sum_{i=1}^{s} b_i^{[\nu]} g^{[\nu]}(A(Y_i)).$$

This shows that $A(y_1) = (A \circ \Phi_f)(y_0) = (\Phi_g \circ A)(y_0)$, so Φ is N-affine equivariant. Finally, if $b_i^{[\nu]} = b_i$ is independent of ν, then Proposition 5.3 and Example 5.4 show that Φ preserves affine invariants. $\square$

Remark 5.6 It is straightforward to show that, in fact, all NB-series methods are N-affine equivariant. (This includes, e.g., generalized additive Runge–Kutta methods,

whose symplecticity conditions were recently investigated by Günther, Sandu, and Zanna [13].) The proof is, essentially, to repeatedly differentiate the A-relatedness condition $A' \circ f^{[\nu]} = g^{[\nu]} \circ A$, obtaining a relation between the elementary differentials.

Theorem 5.7 *Let $\mathcal{F}$ satisfy Assumption 2.8. An additive numerical integrator Φ preserves $\mathcal{F}$-invariants if and only if it is $\mathcal{F}$-functionally equivariant and preserves affine invariants.*

Proof ($\Rightarrow$) Suppose Φ preserves $\mathcal{F}$-invariants. The proof of $\mathcal{F}$-functional equivariance is essentially identical to that in Theorem 2.9, and preservation of affine invariants follows from the fact that $\mathcal{F}$ contains affine maps by Assumption 2.8.

($\Leftarrow$) Conversely, suppose that Φ is $\mathcal{F}$-functionally equivariant and preserves affine invariants. If $F \in \mathcal{F}(Y, Z)$ is an invariant of $f \in \mathfrak{X}(Y)$, then $g^{[\nu]}(y, z) = \big(f^{[\nu]}(y), F'(y) f^{[\nu]}(y) \big)$ is the corresponding decomposition of $g = (f, 0)$. By $\mathcal{F}$-functional equivariance, we have $\Phi_{g^{[1]}, \ldots, g^{[N]}} \colon \big(y_0, F(y_0) \big) \mapsto \big(y_1, F(y_1) \big)$. Finally, since $G(y, z) = z$ is an affine invariant of g, it is preserved by $\Phi_{g^{[1]}, \ldots, g^{[N]}}$, and thus $F(y_0) = F(y_1)$. □

Example 5.8 Let $\mathcal{F}$ be the class of quadratic maps. It follows that an additive numerical integrator preserves quadratic invariants if and only if it is quadratic functionally equivariant and preserves affine invariants. For ARK methods, a sufficient condition is that $b_i^{[\nu]} = b_i$ be independent of ν and $b_i^{[\nu]} a_{ij}^{[\mu]} + b_j^{[\mu]} a_{ji}^{[\nu]} = b_i^{[\nu]} b_j^{[\mu]}$ for all i, j, μ, ν. The proof is identical to that for symplecticity of ARK methods, cf. Aráujo, Murua, and Sanz-Serna [2, Theorem 7].

Splitting methods take $\Phi_{f^{[1]}, \ldots, f^{[N]}}$ to be a composition of exact flows $\varphi_{\tau f^{[\nu]}}$, i.e.,

$$\Phi_{f^{[1]}, \ldots, f^{[N]}} = \varphi_{\tau_s f^{[\nu_s]}} \circ \cdots \circ \varphi_{\tau_1 f^{[\nu_1]}},$$

where consistency requires $\sum_{\nu_i = \nu} \tau_i = 1$ for all $\nu = 1, \ldots, N$. For $N = 2$, the two most elementary splitting methods are the Lie–Trotter splitting $\varphi_{f^{[1]}} \circ \varphi_{f^{[2]}}$ and the Strang splitting $\varphi_{\frac{1}{2} f^{[2]}} \circ \varphi_{f^{[1]}} \circ \varphi_{\frac{1}{2} f^{[2]}}$, where φ denotes the exact time-1 flow. Since the exact flow is equivariant (and hence functionally equivariant) with respect to *all* maps F, the chain rule implies that this is also true of splitting methods. As a consequence of Theorem 5.7, we get the following negative result for splitting methods.

Corollary 5.9 *Any splitting method that preserves affine invariants equals the exact flow.*

Proof Since splitting methods are equivariant with respect to all maps, Theorem 5.7 implies that any splitting method preserving affine invariants preserves *all* invariants. To see that this must be the exact flow, consider the vector field $(f, 1) \in \mathfrak{X}(Y \times \mathbb{R})$, which augments $\dot{y} = f(y)$ by the equation $\dot{t} = 1$. The exact solution is $y(t) = \varphi_{tf}(y_0)$, so $F(y, t) = y - \varphi_{tf}(y_0)$ is an invariant of $(f, 1)$. Therefore, $F(y_1, 1) = F(y_0, 0) = 0$, which says that $y_1 = \varphi_f(y_0)$. □

5.2 Partitioned Methods

We finally consider partitioned methods, which are based on a partitioning $Y = Y^{[1]} \oplus \cdots \oplus Y^{[N]}$. In particular, we have in mind partitioned Runge–Kutta and P-series methods (cf. Hairer [14]). These are closely related to the methods in the previous section, except the vector field decomposition $f = f^{[1]} + \cdots + f^{[N]}$ is uniquely specified by the partitioning of Y, i.e., $f^{[\nu]}(y) \in Y^{[\nu]}$ for all $y \in Y$ and $\nu = 1, \ldots, N$. For this reason, we write the flow of such a method as Φ_f rather than $\Phi_{f^{[1]},\ldots,f^{[N]}}$. By a partitioned numerical integrator, we mean the entire collection of maps $\Phi = \{\Phi_f : f \in \mathfrak{X}(Y),\ Y = \bigoplus_{\nu=1}^N Y^{[\nu]} \text{ a partitioned Banach space}\}$.

Definition 5.10 Given partitioned spaces $Y = \bigoplus_{\nu=1}^N Y^{[\nu]}$ and $U = \bigoplus_{\nu=1}^N U^{[\nu]}$, we say that $A \colon Y \to U$ is a *P-affine map* if it decomposes as $A = \bigoplus_{\nu=1}^N A^{[\nu]}$, where each $A^{[\nu]} \colon Y^{[\nu]} \to U^{[\nu]}$ is affine. A partitioned numerical integrator Φ is *P-affine equivariant* if $A \circ \Phi_f = \Phi_g \circ A$ whenever $f^{[\nu]}$ and $g^{[\nu]}$ are A-related for all $\nu = 1, \ldots, N$, all P-affine maps A, all partitionings, and all Banach spaces Y and U.

Example 5.11 If we partition $U = \mathbb{R}$ into $U^{[\mu]} = \mathbb{R}$ and $U^{[\nu]} = \{0\}$ for $\nu \neq \mu$, then the P-affine functionals are those depending only on $Y^{[\mu]}$. Affine functionals depending on more than one component $Y^{[\nu]}$ cannot be P-affine for any partitioning of $\mathbb{R}$. In particular, if we take $Y = \mathbb{R}^2 = (\mathbb{R} \times \{0\}) \oplus (\{0\} \times \mathbb{R})$, then:

- $(q, p) \mapsto q$ is P-affine for the partitioning $U = \mathbb{R} \oplus \{0\}$;
- $(q, p) \mapsto p$ is P-affine for the partitioning $U = \{0\} \oplus \mathbb{R}$;
- $(q, p) \mapsto q + p$ is never P-affine.

Proposition 5.12 *If an additive numerical integrator Ψ is N-affine equivariant, then the partitioned numerical integrator Φ defined by $\Phi_f = \Psi_{f^{[1]},\ldots,f^{[N]}}$ is P-affine equivariant.*

Proof This follows immediately from the definitions, since P-affine maps are affine. $\square$

Example 5.13 (partitioned Runge–Kutta methods) An s-stage partitioned Runge–Kutta method (PRK) is just the application of an ARK method to a partitioned space, as in Proposition 5.12, where Φ is the PRK method and Ψ is the ARK method. As an immediate corollary of this proposition, all PRK methods are P-affine equivariant.

The definition of F- and $\mathcal{F}$-functional equivariance is the same as in Definition 2.4, where given $Y = \bigoplus_{\nu=1}^N Y^{[\nu]}$ and $Z = \bigoplus_{\nu=1}^N Z^{[\nu]}$, we partition $Y \times Z = \bigoplus_{\nu=1}^N (Y^{[\nu]} \times Z^{[\nu]})$. However, the methods being considered are not necessarily equivariant with respect to all affine maps, so Assumption 2.8 is too restrictive on $\mathcal{F}$. We therefore replace it with the following, which just replaces "affine" by "P-affine" for specified partitions.

Assumption 5.14 Assume that:

- $\mathcal{F}(Y, Y)$ contains the identity map for all $Y = \bigoplus_{\nu=1}^N Y^{[\nu]}$;

- $\mathcal{F}(Y, Z)$ is a vector space for all $Y = \bigoplus_{\nu=1}^{N} Y^{[\nu]}$ and $Z = \bigoplus_{\nu=1}^{N} Z^{[\nu]}$;
- $\mathcal{F}$ is invariant under composition with P-affine maps, in the following sense: If $A \colon Y \to U$ and $B \colon V \to Z$ are P-affine and $F \in \mathcal{F}(U, V)$, then $B \circ F \circ A \in \mathcal{F}(Y, Z)$, for all $Y = \bigoplus_{\nu=1}^{N} Y^{[\nu]}$, $Z = \bigoplus_{\nu=1}^{N} Z^{[\nu]}$, $U = \bigoplus_{\nu=1}^{N} U^{[\nu]}$, and $V = \bigoplus_{\nu=1}^{N} V^{[\nu]}$.

Theorem 5.15 *Let $\mathcal{F}$ satisfy Assumption 5.14. A partitioned numerical integrator Φ preserves $\mathcal{F}$-invariants if and only if it is $\mathcal{F}$-functionally equivariant.*

Proof The proof is formally identical to that for Theorem 2.9.   □

Example 5.16 Let $\mathcal{F}$ be the class of P-affine maps. It follows that all P-affine equivariant methods preserve P-affine invariants. In particular, by Example 5.11, affine invariants $F \colon Y \to \mathbb{R}$ depending only on a single component $Y^{[\mu]}$ are preserved.

Example 5.17 Let $\mathcal{F}$ be the class of *all* affine maps, irrespective of partitioning. It follows that P-affine equivariant methods preserve affine invariants if and only if they are affine functionally equivariant. For PRK methods, as for ARK methods, this holds if $b_i^{[\nu]} = b_i$ is independent of ν. (See Example 5.4 and Proposition 5.5.)

Example 5.18 Let $\mathcal{F}$ be the class of quadratic maps that are at most bilinear with respect to the partition. i.e., terms may be bilinear in $y^{[\mu]}$ and $y^{[\nu]}$ for $\mu \neq \nu$. For PRK methods, a sufficient condition for $\mathcal{F}$-invariant preservation, and thus for $\mathcal{F}$-functional equivariance, is that $b_i^{[\nu]} = b_i$ be independent of ν and $b_i^{[\nu]} a_{ij}^{[\mu]} + b_j^{[\mu]} a_{ji}^{[\nu]} = b_i^{[\nu]} b_j^{[\mu]}$ for all i, j, and $\mu \neq \nu$. This is a straightforward generalization of the $N = 2$ case, cf. Hairer, Lubich, and Wanner [15, Theorem IV.2.4].

Example 5.19 Let $\mathcal{F}$ be the class of *all* quadratic maps, irrespective of partitioning. For PRK methods, as for ARK methods, a sufficient condition for quadratic invariant preservation, and thus for quadratic functional equivariance, is that $b_i^{[\nu]} = b_i$ be independent of ν and $b_i^{[\nu]} a_{ij}^{[\mu]} + b_j^{[\mu]} a_{ji}^{[\nu]} = b_i^{[\nu]} b_j^{[\mu]}$ for all i, j, μ, ν. (See Example 5.8.)

5.3 Closure under Differentiation and (Multi)Symplecticity

Finally, we generalize Theorem 2.12, which allows the functional equivariance results for N-affine and P-affine equivariant methods to be applied to observables depending on variations.

Theorem 5.20 *N-affine and P-affine equivariant methods are closed under differentiation.*

Proof The proof is basically the same as Theorem 2.12, although we need to specify how Φ is applied to the augmented system

$$\dot{x} = f(x), \qquad \dot{y} = f(y), \qquad \dot{z} = \frac{f(x) - f(y)}{\epsilon}.$$

We simply use the same decomposition or partition for each of the three parts. Specifically, if Φ is N-affine equivariant, then we decompose

$$f(x) = \sum_{\nu=1}^{N} f^{[\nu]}(x),$$

$$f(y) = \sum_{\nu=1}^{N} f^{[\nu]}(y),$$

$$\frac{f(x) - f(y)}{\epsilon} = \sum_{\nu=1}^{N} \frac{f^{[\nu]}(x) - f^{[\nu]}(y)}{\epsilon},$$

while if Φ is P-affine equivariant, we partition $Y \times Y \times Y = \bigoplus_{\nu=1}^{N}(Y^{[\nu]} \times Y^{[\nu]} \times Y^{[\nu]})$. The proof then proceeds as in Theorem 2.12. $\qquad\square$

Therefore, the results on symplecticity and multisymplecticity of affine equivariant methods preserving quadratic invariants also hold for N-affine and P-affine equivariant methods preserving quadratic invariants. Moreover, since the canonical symplectic form $\omega = \mathrm{d}q^i \wedge \mathrm{d}p_i$ and multisymplectic form $\omega^0 = \mathrm{d}u^i \wedge \mathrm{d}p_i$ are bilinear on $Y = V \times V^*$, it suffices for an $N = 2$ partitioned method to preserve only bilinear invariants, as in Example 5.18. This includes widely used symplectic PRK methods such as Störmer/Verlet and the Lobatto IIIA–IIIB pair (Hairer, Lubich, and Wanner [15, Sections IV.2 and VI.4]), as well as compositions of these methods.

6 Concluding Remarks

We conclude by posing a natural question for future investigation: *Which numerical integrators are affine functionally equivariant?* Here is a summary of some related results that have been mentioned throughout this paper:

- B-series methods are precisely the *affine equivariant* methods [18], so by Proposition 2.6, they are included among the affine functionally equivariant methods.
- Aromatic B-series methods are precisely the *affine isomorphism equivariant* methods [23]. Since only isomorphsims are considered, the series coefficients may vary depending on dim Y. If the series coefficients *are* constant across dimensions, then the method is affine functionally equivariant, as in Remark 2.7. Conversely, variable-coefficient methods cannot be affine functionally equivariant, since y would then evolve differently between the original and augmented systems.
- As shown in Example 5.17, partitioned methods may also be affine functionally equivariant, e.g., a PRK method with $b_i^{[\nu]} = b_i$ independent of ν. However, such methods are generally not affine isomorphism equivariant, e.g., if $a_{ij}^{[\nu]}$ varies with ν, so affine functional equivariance need not imply affine isomorphism equivariance.

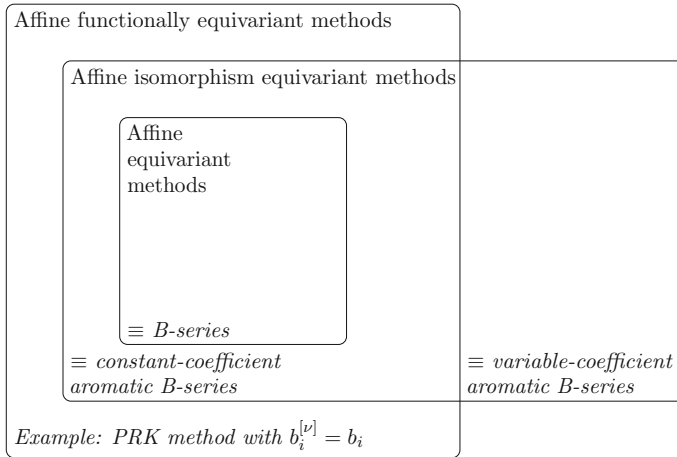Figure 1 depicts these relationships among the different classes of "equivariant" methods.

Fig. 1 The landscape of equivariant methods

References

1. R. Abraham, J. E. Marsden, and T. Ratiu, Manifolds, tensor analysis, and applications, vol. 75 of Applied Mathematical Sciences, Springer-Verlag, New York, second ed., 1988.
2. A. L. Araújo, A. Murua, and J. M. Sanz-Serna, Symplectic methods based on decompositions, SIAM J. Numer. Anal., 34 (1997), pp. 1926–1947.
3. U. M. Ascher and R. I. McLachlan, Multisymplectic box schemes and the Korteweg–de Vries equation, Appl. Numer. Math., 48 (2004), pp. 255–269.
4. Y. Berchenko-Kogan and A. Stern, Constraint-preserving hybrid finite element methods for Maxwell's equations, Found. Comput. Math., 21 (2021), pp. 1075–1098.
5. P. B. Bochev and C. Scovel, On quadratic invariants and symplectic structure, BIT, 34 (1994), pp. 337–345.
6. T. J. Bridges and S. Reich, Multi-symplectic integrators: numerical schemes for Hamiltonian PDEs that conserve symplecticity, Phys. Lett. A, 284 (2001), pp. 184–193.
7. K. Burrage and J. C. Butcher, Stability criteria for implicit Runge-Kutta methods, SIAM J. Numer. Anal., 16 (1979), pp. 46–57.
8. J. C. Butcher, A stability property of implicit Runge–Kutta methods, BIT, 15 (1975), pp. 358–361.
9. P. Chartier and A. Murua, Preserving first integrals and volume forms of additively split systems, IMA J. Numer. Anal., 27 (2007), pp. 381–405.
10. B. Cockburn, J. Gopalakrishnan, and R. Lazarov, Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems, SIAM J. Numer. Anal., 47 (2009), pp. 1319–1365.
11. T. de Donder, Théorie Invariantive du Calcul des Variations, Gauthier-Villars, second ed., 1935.
12. G. Frasca-Caccia and P. E. Hydon, A new technique for preserving conservation laws, Foundations of Computational Mathematics, (2021). https://doi.org/10.1007/s10208-021-09511-1.
13. M. Günther, A. Sandu, and A. Zanna, Symplectic GARK methods for Hamiltonian systems, 2021. Preprint, 2103.04110 [math.NA].

14. E. Hairer, Order conditions for numerical methods for partitioned ordinary differential equations, Numer. Math., 36 (1980/81), pp. 431–445.
15. E. Hairer, C. Lubich, and G. Wanner, Geometric numerical integration, vol. 31 of Springer Series in Computational Mathematics, Springer, Heidelberg, 2010.
16. A. Iserles, G. R. W. Quispel, and P. S. P. Tse, B-series methods cannot be volume-preserving, BIT, 47 (2007), pp. 351–378.
17. R. McLachlan and M. Perlmutter, Conformal Hamiltonian systems, J. Geom. Phys., 39 (2001), pp. 276–300.
18. R. I. McLachlan, K. Modin, H. Munthe-Kaas, and O. Verdier, B-series methods are exactly the affine equivariant methods, Numer. Math., 133 (2016), pp. 599–622.
19. R. I. McLachlan and G. R. W. Quispel, What kinds of dynamics are there? Lie pseudogroups, dynamical systems and geometric integration, Nonlinearity, 14 (2001), pp. 1689–1705.
20. R. I. McLachlan and G. R. W. Quispel, Splitting methods, Acta Numer., 11 (2002), pp. 341–434.
21. R. I. McLachlan, B. N. Ryland, and Y. Sun, High order multisymplectic Runge–Kutta methods, SIAM Journal on Scientific Computing, 36 (2014), pp. A2199–A2226.
22. R. I. McLachlan and A. Stern, Multisymplecticity of hybridizable discontinuous Galerkin methods, Found. Comput. Math., 20 (2020), pp. 35–69.
23. H. Munthe-Kaas and O. Verdier, Aromatic Butcher series, Found. Comput. Math., 16 (2016), pp. 183–215.
24. J.-C. Nédélec, Mixed finite elements in $\mathbb{R}^3$, Numer. Math., 35 (1980), pp. 315–341.
25. P. J. Olver, Applications of Lie groups to differential equations, vol. 107 of Graduate Texts in Mathematics, Springer-Verlag, New York, second ed., 1993.
26. S. Reich, Multi-symplectic Runge-Kutta collocation methods for Hamiltonian wave equations, J. Comput. Phys., 157 (2000), pp. 473–499.
27. B. N. Ryland and R. I. McLachlan, On multisymplecticity of partitioned Runge-Kutta methods, SIAM J. Sci. Comput., 30 (2008), pp. 1318–1340.
28. M. A. Sánchez, C. Ciuca, N. C. Nguyen, J. Peraire, and B. Cockburn, Symplectic Hamiltonian HDG methods for wave propagation phenomena, J. Comput. Phys., 350 (2017), pp. 951–973.
29. Z. Sun and Y. Xing, On structure-preserving discontinuous Galerkin methods for Hamiltonian partial differential equations: energy conservation and multi-symplecticity, J. Comput. Phys., 419 (2020), pp. 109662, 25.
30. H. Weyl, Geodesic fields in the calculus of variation for multiple integrals, Ann. of Math. (2), 36 (1935), pp. 607–629.