



Multisymplecticity of Hybridizable Discontinuous Galerkin Methods

Robert I. McLachlan¹ · Ari Stern²

Received: 24 May 2017 / Revised: 12 September 2018 / Accepted: 22 November 2018 /
Published online: 22 April 2019
© SFoCM 2019

Abstract

In this paper, we prove necessary and sufficient conditions for a hybridizable discontinuous Galerkin method to satisfy a multisymplectic conservation law, when applied to a canonical Hamiltonian system of partial differential equations. We show that these conditions are satisfied by the “hybridized” versions of several of the most commonly used finite element methods, including mixed, nonconforming, and discontinuous Galerkin methods. (Interestingly, for the continuous Galerkin method in dimension greater than one, we show that multisymplecticity only holds in a weaker sense.) Consequently, these general-purpose finite element methods may be used for structure-preserving discretization (or semidiscretization) of canonical Hamiltonian systems of ODEs or PDEs. This establishes multisymplecticity for a large class of arbitrarily high-order methods on unstructured meshes.

Keywords Hybridizable discontinuous Galerkin methods · HDG methods · Multisymplectic methods · Hamiltonian PDEs

Mathematics Subject Classification 65N30 · 37K05

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 5301, New Zealand

² Department of Mathematics and Statistics, Washington University in St. Louis, One Brookings Drive, Campus Box 1146, Saint Louis, MO 63130, USA

1 Introduction

1.1 Motivation and Background

Hamiltonian systems of ordinary differential equations (ODEs) and partial differential equations (PDEs) are ubiquitous in applications, especially in the modeling of physical systems.

One essential property of a Hamiltonian ODE is that its time flow is *symplectic*: that is, it conserves a closed, nondegenerate 2-form on phase space. This has motivated the development of *symplectic integrators*: one-step numerical integrators that, when applied to Hamiltonian ODEs, are also symplectic maps. It turns out that these methods have several numerical advantages that result from preserving the symplectic structure. Furthermore, most of these methods (such as symplectic partitioned Runge–Kutta methods) may be applied to general systems of ODEs, whether or not the user is aware of any Hamiltonian/symplectic structure—but if such a structure is present, then the methods will automatically preserve them. For a comprehensive survey of structure-preserving numerical integrators, including symplectic integrators, see Hairer et al. [22].

Similarly, Hamiltonian PDEs satisfy a *multisymplectic conservation law*. Since symplecticity is a desirable property for numerical integration of canonical Hamiltonian ODEs, it is natural to seek numerical methods for canonical Hamiltonian PDEs whose solutions satisfy the multisymplectic conservation law, in an appropriate sense. There has been important work on multisymplectic methods over the past two decades, particularly by Marsden and collaborators [25, 27–29] and Reich and collaborators [10, 16, 35, 37, 38]. However, most of these methods have consisted either of tensor products of symplectic Runge–Kutta-type methods on rectangular grids or of relatively low-order, finite-difference-type methods on unstructured meshes. For the variational integrators of Marsden et al., one must also know the (Lagrangian) geometric structure of the PDE, in advance, in order to devise the method.

The impact of multisymplecticity on solutions to PDEs, and on their discretizations, is not fully understood. However, multisymplecticity is known to be necessary to preserve traveling waves of hyperbolic equations [30], and compact multisymplectic methods can preserve dispersion relations much better than non-multisymplectic methods or noncompact finite difference methods [5, 33, 34]. For boundary value problems, multisymplecticity restricts the types of bifurcations that can occur [31, 32]. Because it is a local property, multisymplecticity is a strictly stronger property than the symplecticity obtained by integrating over space.

There has been some previous work on the application of finite element methods to certain problems on structured (especially rectangular) meshes. Guo et al. [21] considered the 2-D nonlinear Poisson equation on a regular rectangular grid, meshed with biased triangles, using the continuous Galerkin method with linear shape functions, and they showed that the degrees of freedom satisfied a multisymplectic finite-difference scheme. Zhen et al. [44] did the same for first-order rectangular Lagrange elements on a regular grid. Chen [11] used first- and second-order rectangular elements to derive Lagrangian variational methods (in the sense of Marsden et al. [27]) and applied these

to the sine-Gordon equation on a regular grid, while pointing out that higher-order elements could also be used in principle. In all of these examples, however, finite elements were really only used as a tool to construct a finite-difference stencil on a regular, 2-D rectangular grid.

As finite element methods are traditionally formulated, there is a serious conceptual obstacle to discussing multisymplecticity. Namely, many classical finite element methods are posed on spaces of *global* functions, making it difficult to make sense of *local* properties like the multisymplectic conservation law. Indeed, the interpretation of multisymplecticity is much more straightforward for finite difference or finite volume methods with local stencils, or for tensor products of 1-D integrators on a rectangular grid, which may explain why the previous work has been focused on such methods.

Hybrid finite element methods provide a way around this obstacle, since they consist of *local problems coupled through their boundary traces*, where the boundary traces are allowed to be independent variables. (Oftentimes, these boundary traces are interpreted as “Lagrange multipliers” enforcing weak continuity between local regions.) While hybrid methods have a long history (see the comprehensive work by Brezzi and Fortin [7]), the recent work of Cockburn et al. [13] has shown that a wide variety of finite element methods—including those not previously thought of as hybrid methods—may be “hybridized” within a unified framework. Such methods are called hybridizable discontinuous Galerkin (HDG) methods, and they include not only several classical mixed and hybrid methods, but also hybridized versions of continuous and discontinuous Galerkin methods, nonconforming methods, and others. This framework provides precisely the local structure needed to examine multisymplecticity of the finite element methods in this class.

1.2 Organization of the Paper

The paper is organized as follows:

- In Sect. 2, we review systems of PDEs in a particular canonical form. This form includes the de Donder–Weyl equations for a Hamiltonian and many elliptic and hyperbolic variational PDEs. We recall the multisymplectic conservation law for classical (i.e., smooth) solutions and illustrate how this manifests concretely for a class of semilinear elliptic PDEs in mixed form. We also discuss the relationship between multisymplecticity of solutions and reciprocity principles in physical systems.
- In Sect. 3, we develop a hybrid “flux formulation” for these systems of PDEs. As in Cockburn et al. [13], this yields a collection of weak problems on non-overlapping subdomains, coupled only through approximate traces on their shared boundaries. Our framework includes not only the linear second-order elliptic PDEs considered by Cockburn et al. [13], but also a more general class of nonlinear systems of PDEs, including canonical Hamiltonian PDEs.

Within this formulation, we establish criteria for solutions to satisfy weak and strong versions of the multisymplectic conservation law; the distinction is shown to be related to weak and strong conservativity of numerical fluxes (cf. Arnold et

al. [4], Cockburn et al. [13]). In addition to the subsequent numerical applications, we also use a domain-decomposition argument to write the weak problem (in the sense of distributions) in this flux formulation, thereby establishing multisymplecticity for weak solutions to Hamiltonian PDEs.

- In Sect. 4, we examine several particular classes of HDG methods, including the hybridized Raviart–Thomas (RT-H), Brezzi–Douglas–Marini (BDM-H), local discontinuous Galerkin (LDG-H), continuous Galerkin (CG-H), nonconforming (NC-H), and interior penalty (IP-H) methods. These methods are posed in the framework of Sect. 3, and their multisymplecticity is then examined. Each of these methods, except for CG-H, is proved to be strongly multisymplectic. A counterexample shows that CG-H is only weakly multisymplectic, resulting from the fact that it is only weakly conservative.

2 Canonical and Multisymplectic Systems of PDEs

2.1 Canonical Systems of PDEs

Given a domain $U \subset \mathbb{R}^m$, consider a system of first-order PDEs having the form

$$\partial_\mu u^i = \phi_\mu^i(\cdot, u, \sigma), \quad -\partial_\mu \sigma_i^\mu = f_i(\cdot, u, \sigma), \quad (1)$$

where $\mu = 1, \dots, m$ and $i = 1, \dots, n$. Here, $u = u^i(x)$ and $\sigma = \sigma_i^\mu(x)$ are unknown functions on U , while $\phi = \phi_\mu^i(x, u, \sigma)$ and $f = f_i(x, u, \sigma)$ are given functions on $U \times \mathbb{R}^n \times \mathbb{R}^{mn}$. We abbreviate $\partial_\mu := \partial/\partial x^\mu$ and adopt the Einstein index convention of summing over repeated indices—so, for instance, the expression $\partial_\mu \sigma_i^\mu$ in (1) has an implied sum over μ and may therefore be interpreted as the divergence of σ_i .

Among these is the important class of (*canonical*) *Hamiltonian systems*,

$$\partial_\mu u^i = \frac{\partial H}{\partial \sigma_i^\mu}, \quad -\partial_\mu \sigma_i^\mu = \frac{\partial H}{\partial u^i}. \quad (2)$$

where $H = H(x, u, \sigma)$ is a function called the *Hamiltonian*. In the special case $m = 1$, the resulting system of ODEs yields Hamilton's equations of classical mechanics, which are usually written as

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \quad -\dot{p}_i = \frac{\partial H}{\partial q^i}.$$

The equations in (2) are called the *de Donder–Weyl equations* (de Donder [15] and Weyl [43]). These canonical systems are an important special case of a more general class of Hamiltonian systems of PDEs, cf. Bridges [8,9].

Throughout this section, we assume that all of the functions above are smooth. Later, in Sect. 3, we will relax this assumption in order to introduce a weak formulation of (1).

Example 1 (Semilinear Elliptic PDE) Let $n = 1$, so that $u = u(x)$ is a scalar field and $\sigma = \sigma^\mu(x)$ a vector field on $U \subset \mathbb{R}^m$. Consider

$$H(x, u, \sigma) = \frac{1}{2} a_{\mu\nu}(x) \sigma^\mu \sigma^\nu + F(x, u),$$

where $a = a^{\mu\nu}(x)$ is symmetric and positive-definite with matrix inverse $a_{\mu\nu}(x) := (a^{\mu\nu}(x))^{-1}$ at each $x \in U$, and where F is arbitrary. Then the de Donder–Weyl equations for this Hamiltonian are

$$\partial_\mu u = a_{\mu\nu} \sigma^\nu, \quad -\partial_\mu \sigma^\mu = \frac{\partial F}{\partial u}.$$

From the first of these equations, we have $\sigma^\mu = a^{\mu\nu} \partial_\nu u$, so substituting this into the second yields

$$-\partial_\mu a^{\mu\nu} \partial_\nu u = \frac{\partial F}{\partial u},$$

which is a second-order semilinear elliptic PDE in divergence form. Note that we could also have written the de Donder–Weyl equations in the equivalent, coordinate-free form,

$$\text{grad } u = a^{-1} \sigma, \quad -\text{div } \sigma = \frac{\partial F}{\partial u},$$

so the substitution $\sigma = a \text{ grad } u$ yields

$$-\text{div}(a \text{ grad } u) = \frac{\partial F}{\partial u},$$

which is an equivalent expression for the second-order PDE above.

An important special case is when $F(x, u) = f(x)u - \frac{1}{2}c(x)u^2$ for given f and c on U . In this case, we obtain a linear second-order elliptic PDE,

$$-\text{div}(a \text{ grad } u) + cu = -\partial_\mu a^{\mu\nu} \partial_\nu u + cu = f.$$

In particular, if $a^{\mu\nu} \equiv \delta^{\mu\nu}$ (where δ is the Kronecker delta, i.e., a is the identity matrix) and $c \equiv 0$, then this simply becomes Poisson’s equation $-\Delta u = f$ on U .

We will regularly return to this example throughout the paper.

2.2 The Multisymplectic Conservation Law

Define the collection of *canonical 2-forms* $\omega^\mu := du^i \wedge d\sigma_i^\mu$ on $\mathbb{R}^n \times \mathbb{R}^{mn}$, for $\mu = 1, \dots, m$.

Notation Unless otherwise stated, differential forms and exterior differential operators (such as d , $\wedge$, etc.) are assumed to be on $\mathbb{R}^n \times \mathbb{R}^{mn}$, where $x \in U$ (if it appears) is fixed. Differentiation with respect to x will always be denoted using the previously defined ∂_μ notation.

Definition 1 Let (u, σ) be a solution to (1). A (first) variation of (u, σ) is a solution (v, τ) of the linearized problem

$$\partial_\mu v^i = \frac{\partial \phi_\mu^i}{\partial u^j}(\cdot, u, \sigma) v^j + \frac{\partial \phi_\mu^i}{\partial \sigma_j^\nu}(\cdot, u, \sigma) \tau_j^\nu, \quad (3a)$$

$$-\partial_\mu \tau_i^\mu = \frac{\partial f_i}{\partial u^j}(\cdot, u, \sigma) v^j + \frac{\partial f_i}{\partial \sigma_j^\nu}(\cdot, u, \sigma) \tau_j^\nu. \quad (3b)$$

The system (1) is *multisymplectic* if $\partial_\mu \left(\omega^\mu((v, \tau), (v', \tau')) \right) = 0$ for any pair of variations (v, τ) and (v', τ') . This is abbreviated by

$$\partial_\mu \omega^\mu = 0, \quad (4)$$

where it is understood that ω^μ is evaluated on variations of solutions to (1). Equation (4) is called the *multisymplectic conservation law*.

Lemma 1 The system (1) is multisymplectic if and only if the 1-form $\phi_\mu^i d\sigma_i^\mu + f_i du^i$ on $\mathbb{R}^n \times \mathbb{R}^{mn}$ is closed for each $x \in U$.

Proof Using (1), we calculate

$$\begin{aligned} \partial_\mu \omega^\mu &= d(\partial_\mu u^i) \wedge d\sigma_i^\mu + du^i \wedge d(\partial_\mu \sigma_i^\mu) \\ &= d\phi_\mu^i \wedge d\sigma_i^\mu - du^i \wedge df_i \\ &= d(\phi_\mu^i d\sigma_i^\mu + f_i du^i), \end{aligned}$$

so the first expression vanishes if and only if the last expression vanishes. $\square$

Remark 1 Certain steps in this calculation, such as $\partial_\mu(du^i) = d(\partial_\mu u^i)$, are seen to be valid by evaluating both sides on arbitrary variations of (u, σ) :

$$\partial_\mu(du^i(v, \tau)) = \partial_\mu v^i = d\phi_\mu^i(v, \tau) = d(\partial_\mu u)(v, \tau).$$

Similar calculations with differential forms will appear throughout this paper, where they are interpreted as holding for arbitrary variations of solutions.

Corollary 1 Every Hamiltonian system is multisymplectic.

Proof From (2), we have

$$\phi_\mu^i d\sigma_i^\mu + f_i du^i = \frac{\partial H}{\partial \sigma_i^\mu} d\sigma_i^\mu + \frac{\partial H}{\partial u^i} du^i = dH,$$

which is exact and therefore closed. $\square$

Corollary 2 Every multisymplectic system is Hamiltonian.

Proof Since $\mathbb{R}^n \times \mathbb{R}^{mn}$ is simply connected, its first de Rham cohomology is trivial. Hence, the closed 1-form $\phi_\mu^i d\sigma_i^\mu + f_i du^i$ is exact, i.e., it equals dH for some Hamiltonian H . More precisely, for each fixed $x \in U$, the 1-form equals dH_x for some function H_x on $\mathbb{R}^n \times \mathbb{R}^{mn}$, and these may be combined into a single Hamiltonian $H(x, u, \sigma) := H_x(u, \sigma)$. $\square$

Remark 2 Corollary 2 depends entirely on the fact that $\mathbb{R}^n \times \mathbb{R}^{mn}$ has trivial first de Rham cohomology. However, it is possible to define canonical Hamiltonian systems on more general spaces—in particular, on the dual jet bundle of some fiber bundle over U (cf. Gotay [17] and references therein). In this setting, the argument of Corollary 2 holds only if the fibers of this bundle have trivial first de Rham cohomology. However, a weaker statement—that every multisymplectic system is *locally* Hamiltonian—still holds, by Poincaré’s lemma. By contrast, Corollary 1 remains true even in this more general setting.

Example 2 (Semilinear Elliptic PDE, Continued) Let us see how the multisymplectic conservation law manifests in the class of semilinear elliptic PDEs we encountered in Example 1. For the system

$$\partial_\mu u = a_{\mu\nu} \sigma^\nu, \quad -\partial_\mu \sigma^\mu = \frac{\partial F}{\partial u},$$

we calculate

$$\begin{aligned} \partial_\mu \omega^\mu &= \partial_\mu (du \wedge d\sigma^\mu) \\ &= d(\partial_\mu u) \wedge d\sigma^\mu + du \wedge d(\partial_\mu \sigma^\mu) \\ &= a_{\mu\nu} d\sigma^\nu \wedge d\sigma^\mu + du \wedge \left(-\frac{\partial^2 F}{\partial u^2} du - \frac{\partial^2 F}{\partial \sigma \partial u} d\sigma \right). \end{aligned}$$

The first term vanishes by the symmetry of a and the antisymmetry of $\wedge$, while the remaining terms vanish since $du \wedge du = 0$ (again, the antisymmetry of $\wedge$) and since $F = F(x, u)$ does not depend on σ .

2.3 Integral form of the Multisymplectic Conservation Law

Given an arbitrary subdomain $K \subseteq U$, the divergence theorem implies that

$$\int_K \partial_\mu \omega^\mu d^m x = \int_{\partial K} \omega^\mu d^{m-1} x_\mu,$$

where $d^m x := dx^1 \wedge \cdots \wedge dx^m$ is the standard Euclidean volume form on U and $d^{m-1} x_\mu := \iota_{\partial/\partial x^\mu} (d^m x)$, where ι is the interior product (or contraction). Therefore, an equivalent formulation of the multisymplectic conservation law (4) is that

$$\int_{\partial K} \omega^\mu d^{m-1} x_\mu = 0, \quad \forall K \subseteq U. \quad (5)$$

We call this the *integral form of the multisymplectic conservation law*. As with (4), this is interpreted as holding when ω^μ is evaluated on arbitrary variations of a solution to (1).

Note that, by the definition of the $\wedge$ product,

$$\omega^\mu = du^i \wedge d\sigma_i^\mu = du^i \otimes d\sigma_i^\mu - d\sigma_i^\mu \otimes du^i,$$

so (5) may also be written as

$$\int_{\partial K} (du^i \otimes d\sigma_i^\mu) d^{m-1}x_\mu = \int_{\partial K} (d\sigma_i^\mu \otimes du^i) d^{m-1}x_\mu. \quad (6)$$

Hence, the multisymplectic conservation law may be interpreted as a symmetry condition on the Poincaré–Steklov operator mapping Dirichlet boundary conditions for u to the corresponding boundary conditions for σ . See Agoshkov [3], where the symmetry of the Poincaré–Steklov operator is discussed in the context of domain-decomposition methods for linear elliptic problems. Similarly, Belishev and Sharafutdinov [6] establish the symmetry of a Poincaré–Steklov operator for harmonic differential forms on a manifold with boundary.

Example 3 (Semilinear Elliptic PDE, Continued) Let us revisit the class of semilinear elliptic PDEs we encountered in Examples 1 and 2. Let (u, σ) be a solution, and consider the linearized problem

$$\partial_\mu v = a_{\mu\nu} \tau^\nu, \quad -\partial_\mu \tau^\mu = \frac{\partial^2 F}{\partial u^2} v.$$

Here, $\partial^2 F / \partial u^2$ is evaluated at $(x, u(x))$ and hence is a function of x alone. If (v, τ) and (v', τ') are two arbitrary solutions to this problem, then

$$\begin{aligned} \int_{\partial K} v \tau'^\mu d^{m-1}x_\mu &= \int_K [(\partial_\mu v) \tau'^\mu + v(\partial_\mu \tau'^\mu)] d^m x \\ &= \int_K \left(a_{\mu\nu} \tau^\nu \tau'^\mu - v \frac{\partial^2 F}{\partial u^2} v' \right) d^m x. \end{aligned}$$

By a similar calculation, switching (v, τ) with (v', τ') ,

$$\int_{\partial K} v' \tau^\mu d^{m-1}x_\mu = \int_K \left(a_{\mu\nu} \tau'^\nu \tau^\mu - v' \frac{\partial^2 F}{\partial u^2} v \right) d^m x.$$

However, since a is symmetric, these integrals are identical, and we conclude

$$\int_{\partial K} v \tau'^\mu d^{m-1}x_\mu = \int_{\partial K} v' \tau^\mu d^{m-1}x_\mu.$$

That this equality holds for every (v, τ) and (v', τ') is precisely the statement (6) of the multisymplectic conservation law.

In the special case $a^{\mu\nu} \equiv \delta^{\mu\nu}$, we have $\tau = \text{grad } v$ and $\tau' = \text{grad } v'$, so this can be written as

$$\int_{\partial K} v \text{grad } v' \cdot \mathbf{n} = \int_{\partial K} v' \text{grad } v \cdot \mathbf{n}$$

where $\mathbf{n}$ denotes the outer unit normal to ∂K . Hence, in this case, the multisymplectic conservation law expresses the symmetry of the Dirichlet–Neumann operator $v|_{\partial K} \mapsto \text{grad } v \cdot \mathbf{n}|_{\partial K}$ for the linearized problem.

2.4 Multisymplecticity and Reciprocity

In many physical systems, the multisymplectic conservation law is closely tied to so-called *reciprocity* phenomena, such as Green’s reciprocity in electrostatics and Betti reciprocity in elasticity. (See, for example, Abraham and Marsden [1, Section 5.3], Marsden and Hughes [26, Section 5.6], and Lew et al. [25].) These reciprocity phenomena are also exploited, numerically, in formulations of the boundary element method (cf. Partridge et al. [36]). We now briefly discuss the relationship between multisymplecticity and reciprocity, using the language we have developed throughout this section.

Let (u, σ) be a solution to (1). The multisymplectic conservation law (4) is just the statement that

$$\partial_\mu(v^i \tau_i'^\mu) = \partial_\mu(v'^i \tau_i^\mu),$$

where (v, τ) and (v', τ') are arbitrary variations of (u, σ) , i.e., solutions to the linearized problem (3). Integrating both sides over $K \subseteq U$ and applying the divergence theorem gives

$$\int_{\partial K} v^i \tau_i'^\mu \, d^{m-1}x_\mu = \int_{\partial K} v'^i \tau_i^\mu \, d^{m-1}x_\mu,$$

which is the statement of (5) and (6).

We now generalize the above to the case where the variations (v, τ) and (v', τ') each solve a *perturbed* version of the linearized problem.

Definition 2 Given a solution (u, σ) to (1), we say that (v, τ) solves the *linearized problem with incremental sources* ψ and g if

$$\partial_\mu v^i = \frac{\partial \phi_\mu^i}{\partial u^j}(\cdot, u, \sigma) v^j + \frac{\partial \phi_\mu^i}{\partial \sigma_j^v}(\cdot, u, \sigma) \tau_j^v + \psi_\mu^i(\cdot, v, \tau), \quad (7a)$$

$$-\partial_\mu \tau_i^\mu = \frac{\partial f_i}{\partial u^j}(\cdot, u, \sigma) v^j + \frac{\partial f_i}{\partial \sigma_j^v}(\cdot, u, \sigma) \tau_j^v + g_i(\cdot, v, \tau), \quad (7b)$$

where $\psi = \psi_\mu^i(x, v, \tau)$ and $g = g_i(x, v, \tau)$ are given functions.

Let (v, τ) be a solution to the linearized problem with incremental sources ψ and g , and let (v', τ') be a solution to the linearized problem with incremental sources ψ'

and g' . By the Leibniz rule,

$$\begin{aligned}\partial_\mu(v^i \tau_i'^\mu) &= (\partial_\mu v^i) \tau_i'^\mu + v^i (\partial_\mu \tau_i'^\mu), \\ \partial_\mu(v'^i \tau_i^\mu) &= (\partial_\mu v'^i) \tau_i^\mu + v'^i (\partial_\mu \tau_i^\mu),\end{aligned}$$

which we subtract and rearrange to obtain

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

Assuming that (1) is multisymplectic, the ϕ and f terms cancel when we substitute (7), leaving the equation

$$\begin{aligned}\partial_\mu(v^i \tau_i'^\mu) - \psi_\mu^i(\cdot, v, \tau) \tau_i'^\mu + v^i g'_i(\cdot, v', \tau') \\ = \partial_\mu(v'^i \tau_i^\mu) - \psi_\mu^i(\cdot, v', \tau') \tau_i^\mu + v'^i g_i(\cdot, v, \tau).\end{aligned}\quad (8)$$

Integrating over $K \subseteq U$ and applying the divergence theorem gives

$$\begin{aligned}\int_{\partial K} v^i \tau_i'^\mu d^{m-1} x_\mu - \int_K [\psi_\mu^i(\cdot, v, \tau) \tau_i'^\mu - v^i g'_i(\cdot, v', \tau')] d^m x \\ = \int_{\partial K} v'^i \tau_i^\mu d^{m-1} x_\mu - \int_K [\psi_\mu^i(\cdot, v', \tau') \tau_i^\mu + v'^i g_i(\cdot, v, \tau)] d^m x.\end{aligned}\quad (9)$$

Equations (8) and (9) are the differential and integral forms, respectively, of the reciprocity law for a multisymplectic system of PDEs. They may be interpreted as describing a symmetric (or “reciprocal”) relationship between the perturbation of the system by incremental sources and the incremental response of the system to such perturbations. In the special case where the incremental sources vanish, we recover the multisymplectic conservation law.

Example 4 (Semilinear Elliptic PDE, Continued) Let us once again examine the semilinear elliptic PDEs considered in Examples 1, 2, and 3. The linearized problem with incremental sources is

$$\partial_\mu v = a_{\mu\nu} \tau^\nu + \psi_\mu(\cdot, v, \tau), \quad -\partial_\mu \tau^\mu = \frac{\partial^2 F}{\partial u^2} v + g(\cdot, v, \tau).$$

To see how reciprocity arises, we compute

$$\begin{aligned}\partial_\mu(v \tau^\mu) &= [a_{\mu\nu} \tau^\nu + \psi_\mu(\cdot, v, \tau)] \tau'^\mu - v \left[\frac{\partial^2 F}{\partial u^2} v' + g'(\cdot, v', \tau') \right], \\ \partial_\mu(v' \tau^\mu) &= [a_{\mu\nu} \tau'^\nu + \psi'_\mu(\cdot, v', \tau')] \tau^\mu - v' \left[\frac{\partial^2 F}{\partial u^2} v + g(\cdot, v, \tau) \right].\end{aligned}$$

Subtracting, the terms involving a and F cancel by symmetry, yielding

$$\partial_\mu(v \tau'^\mu) - \psi_\mu(\cdot, v, \tau) \tau'^\mu + v g'(\cdot, v', \tau') = \partial_\mu(v' \tau^\mu) - \psi'_\mu(\cdot, v', \tau') \tau^\mu + v' g(\cdot, v, \tau),$$

which is precisely the statement (8). Integrating this over $K \Subset U$ gives

$$\begin{aligned} & \int_{\partial K} v \tau'^\mu \, d^{m-1}x_\mu - \int_{\partial K} [\psi_\mu(\cdot, v, \tau) \tau'^\mu - v g'(\cdot, v', \tau')] \, d^m x \\ &= \int_{\partial K} v' \tau^\mu \, d^{m-1}x_\mu - \int_K [\psi'_\mu(\cdot, v', \tau') \tau^\mu - v' g(\cdot, v, \tau)] \, d^{m-1}x_\mu, \end{aligned}$$

which is the statement (9).

As an important special case, which arises in the primal (or Lagrangian) formulation of this system, suppose that $\psi = 0$ and that $g = g(\cdot, v)$, so that $\tau = a \operatorname{grad} v$ and v solves the second-order equation

$$-\operatorname{div}(a \operatorname{grad} v) = \frac{\partial^2 F}{\partial u^2} v + g(\cdot, v).$$

If the corresponding properties hold for (v', τ') , then we can write the reciprocity law as

$$\operatorname{div}(v \tau') + v g'(\cdot, v') = \operatorname{div}(v' \tau) + v' g(\cdot, v),$$

whose integral form on $K \Subset U$ is

$$\int_{\partial K} v \tau' \cdot \mathbf{n} + \int_K v g'(\cdot, v') = \int_{\partial K} v' \tau \cdot \mathbf{n} + \int_K v' g(\cdot, v),$$

As a final specialization, let $a^{\mu\nu} \equiv \delta^{\mu\nu}$ and $F(x, u) = f(x)u$, so that u satisfies Poisson's equation, $-\Delta u = f$, and v and v' satisfy

$$-\Delta v = g(\cdot, v), \quad -\Delta v' = g'(\cdot, v').$$

Then the reciprocity law, in differential form, is

$$\operatorname{div}(v \tau') - v \Delta v' = \operatorname{div}(v' \tau) - v' \Delta v,$$

while the integral form on $K \Subset U$ is

$$\int_{\partial K} v \tau' \cdot \mathbf{n} - \int_K v \Delta v' = \int_{\partial K} v' \tau \cdot \mathbf{n} - \int_K v' \Delta v.$$

These last two expressions are two of Green's identities from vector calculus. If v and v' are interpreted as scalar potentials for the electrostatic fields τ and τ' , respectively, then this corresponds to Green's reciprocity.

3 The Flux Formulation and Multisymplecticity

3.1 Domain Decomposition and the Flux Formulation

In this section, we introduce a weak formulation of the problem (1), called the *flux formulation*. This decomposes the problem on U into a collection of local solvers for (u, σ) , coupled through the approximate boundary traces $(\widehat{u}, \widehat{\sigma})$. This forms the foundation of the HDG framework of Cockburn et al. [13] and is closely related to the unified DG framework of Arnold et al. [4].

We mention that our presentation of the flux formulation, and of HDG methods, differs from that in Cockburn et al. [13] in a few ways. In particular, Cockburn et al. focus on linear elliptic problems, which allows them to make substantial use of the solution theory for such problems, including well-posedness of the local and global solvers. By contrast, we are interested in obtaining multisymplecticity criteria for the much more general class of systems (1), without assuming anything about the properties of solutions, even their existence and/or uniqueness.

In this section, the function spaces appearing in the flux formulation may be either infinite-dimensional Hilbert spaces or finite-dimensional subspaces (e.g., polynomials up to some degree). This will allow us to prove general multisymplecticity results that apply both to the original, infinite-dimensional problem (Sect. 3.4) and to finite element approximation via HDG methods (Sect. 4).

To begin, observe that if (u, σ) is a smooth solution to (1) on U , and if (v, τ) are arbitrary smooth test functions, then

$$\begin{aligned} \int_U \partial_\mu u^i \tau_i^\mu \, d^m x &= \int_U \phi_\mu^i \tau_i^\mu \, d^m x, \\ - \int_U \partial_\mu \sigma_i^\mu v^i \, d^m x &= \int_U f_i v^i \, d^m x. \end{aligned}$$

If $\mathcal{T}_h$ is a partition of U into non-overlapping domains $K \in \mathcal{T}_h$, then breaking each of the integrals above into a sum over $K \in \mathcal{T}_h$ and integrating by parts,

$$\sum_{K \in \mathcal{T}_h} \int_{\partial K} u^i \tau_i^\mu \, d^{m-1} x_\mu = \sum_{K \in \mathcal{T}_h} \int_K (u^i \partial_\mu \tau_i^\mu + \phi_\mu^i \tau_i^\mu) \, d^m x, \quad (10a)$$

$$\sum_{K \in \mathcal{T}_h} \int_{\partial K} \sigma_i^\mu v^i \, d^{m-1} x_\mu = \sum_{K \in \mathcal{T}_h} \int_K (\sigma_i^\mu \partial_\mu v^i - f_i v^i) \, d^m x. \quad (10b)$$

In a typical finite element application, U will be polyhedral, and $\mathcal{T}_h$ will be a triangulation of U into simplices $K \in \mathcal{T}_h$.

Following Cockburn et al. [13] (as well as Arnold et al. [4]), we now relax the regularity and inter-element continuity assumptions on u, v and σ, τ in (10), and we replace the boundary traces of u and σ on ∂K by approximate traces $\widehat{u}$ and $\widehat{\sigma}$. Specifically, let

$$V(K) \subset [H^2(K)]^n, \quad \Sigma(K) \subset [H^1(K)]^{mn},$$

be specified local function spaces on each $K \in \mathcal{T}_h$, and define discontinuous global spaces on U by

$$V := \{v \in [L^2(U)]^n : v|_K \in V(K), \forall K \in \mathcal{T}_h\} = \prod_{K \in \mathcal{T}_h} V(K),$$

$$\Sigma := \{\tau \in [L^2(U)]^{mn} : \tau|_K \in \Sigma(K), \forall K \in \mathcal{T}_h\} = \prod_{K \in \mathcal{T}_h} \Sigma(K).$$

Next, specify a space of approximate traces of functions in V ,

$$\widehat{V} \subset [L^2(\mathcal{E}_h)]^n,$$

where $\mathcal{E}_h := \bigcup_{K \in \mathcal{T}_h} \partial K$, along with the subspace

$$\widehat{V}_0 := \{\widehat{v} \in \widehat{V} : \widehat{v}|_{\partial U} = 0\}$$

of approximate traces vanishing on the domain boundary ∂U .

The final ingredient in the HDG framework is the *numerical flux* $\widehat{\sigma}$, which we define slightly differently to Cockburn et al. [13]. As mentioned in the introduction to this section, our treatment is equivalent to [13] for linear problems, but it extends more naturally to nonlinear problems, even without assuming existence and uniqueness of solutions. Let

$$\widehat{\Sigma}(\partial K) \subset [L^2(\partial K)]^{mn}$$

be some space of boundary fluxes on ∂K , and define the space of restricted traces $\widehat{V}(\partial K) := \{\widehat{v}|_{\partial K} : \widehat{v} \in \widehat{V}\}$.

Definition 3 A *local flux function* on $K \in \mathcal{T}_h$ is a bounded linear map

$$\Phi_K : V(K) \times \Sigma(K) \times \widehat{V}(\partial K) \times \widehat{\Sigma}(\partial K) \rightarrow [L^2(\partial K)]^{mn}.$$

Denoting $\widehat{\Sigma} := \prod_{K \in \mathcal{T}_h} \widehat{\Sigma}(\partial K) \subset \prod_{K \in \mathcal{T}_h} [L^2(\partial K)]^{mn}$, this extends naturally to a *global flux function*,

$$\Phi : V \times \Sigma \times \widehat{V} \times \widehat{\Sigma} \rightarrow \prod_{K \in \mathcal{T}_h} [L^2(\partial K)]^{mn}.$$

Remark 3 Elements of $\widehat{\Sigma}$ and $\prod_{K \in \mathcal{T}_h} [L^2(\partial K)]^{mn}$ may be interpreted as functions that are double-valued on internal facets of $\mathcal{T}_h$ and single-valued on boundary facets in ∂U .

We now seek solutions $(u, \sigma, \widehat{u}, \widehat{\sigma}) \in V \times \Sigma \times \widehat{V} \times \widehat{\Sigma}$ satisfying

$$\int_{\partial K} \widehat{u}^i \tau_i^\mu d^{m-1}x_\mu = \int_K (u^i \partial_\mu \tau_i^\mu + \phi_\mu^i \tau_i^\mu) d^m x, \quad \forall \tau \in \Sigma(K), \quad (11a)$$

$$\int_{\partial K} \widehat{\sigma}_i^\mu v^i d^{m-1}x_\mu = \int_K (\sigma_i^\mu \partial_\mu v^i - f_i v^i) d^m x, \quad \forall v \in V(K), \quad (11b)$$

$$\int_{\partial K} \Phi_\mu^i(u, \sigma, \widehat{u}, \widehat{\sigma}) \widehat{\tau}_i^\mu \mathbf{n}^v d^{m-1}x_v = 0, \quad \forall \widehat{\tau} \in \widehat{\Sigma}(\partial K), \quad (11c)$$

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

$$\sum_{K \in \mathcal{T}_h} \int_{\partial K} \widehat{\sigma}_i^\mu \widehat{v}^i d^{m-1}x_\mu = 0, \quad \forall \widehat{v} \in \widehat{V}_0, \quad (11d)$$

the latter of which states that the normal component of $\widehat{\sigma}$ is single-valued (at least in a weak sense) on the internal facets of $\mathcal{T}_h$.

In the language of Cockburn et al. [13], we have “local solvers” (11a)–(11b) on each $K \in \mathcal{T}_h$, and these are coupled globally through the numerical flux $\widehat{\sigma}$ by the conservativity condition (11d). A notable distinction between our approach and that of Cockburn et al. [13] is that they assume $\widehat{\sigma} = \widehat{\sigma}(u, \sigma, \widehat{u})$ is a given function of u, σ , and $\widehat{u}$ on each $K \in \mathcal{T}_h$, whereas we define it through the flux functions Φ_K by adding (11c) to the flux formulation.

Definition 4 The *flux formulation* of (1) on $\mathcal{T}_h$ is given by (11), along with choices of the global function space $\widehat{V}$ and, for each $K \in \mathcal{T}_h$, the local function spaces $V(K)$, $\Sigma(K)$, $\widehat{\Sigma}(\partial K)$ and the flux function Φ_K . We call this a *hybridizable discontinuous Galerkin (HDG) method* whenever $\widehat{V}_0$, V , and Σ (but not necessarily $\widehat{V}$ or $\widehat{\Sigma}$) are finite dimensional.

Remark 4 Note that (11) does not impose any particular boundary conditions on $\widehat{u}|_{\partial U}$ or $\widehat{\sigma}|_{\partial U}$. Hence, (11) corresponds to the system of PDEs (1) rather than a particular boundary value problem associated with (1).

We remain agnostic about the choice of boundary conditions for two reasons. First, multisymplecticity is not a statement about a particular solution, but a statement about variations within a general family of solutions. If we pick out an isolated solution (e.g., by the imposition of boundary conditions), then the “family” of solutions becomes zero-dimensional, so any statement about variations is vacuous. Second, the class of PDEs (1) is quite general, including both elliptic and hyperbolic PDEs, among others, depending on ϕ and f . In the hyperbolic case, when U is a spacetime region, we are not free to impose Dirichlet conditions on all of ∂U .

3.2 Local Multisymplecticity Criteria

In the context of smooth solutions to (1), where ϕ_μ^i and f_i are smooth functions on $U \times \mathbb{R}^n \times \mathbb{R}^{mn}$, Lemma 1 states that multisymplecticity holds if and only if the smooth 1-form $\phi_\mu^i d\sigma_i^\mu + f_i du^i$ is closed for all $x \in U$. For the flux formulation (11), we wish

to relax these smoothness assumptions and express the multisymplecticity condition in terms of function spaces, rather than in a pointwise sense at each $x \in U$.

Observe that (11a)–(11b) still makes sense even if we only have $\phi_\mu^i, f_i \in L^2(U)$ for $\mu = 1, \dots, m$ and $i = 1, \dots, n$. Therefore, rather than assuming that ϕ and f are smooth, let us assume only that

$$\phi_K: V(K) \times \Sigma(K) \rightarrow [L^2(K)]^{mn}, \quad f_K: V(K) \times \Sigma(K) \rightarrow [L^2(K)]^n,$$

for each $K \in \mathcal{T}_h$, which may be extended naturally to

$$\phi: V \times \Sigma \rightarrow [L^2(U)]^{mn}, \quad f: V \times \Sigma \rightarrow [L^2(U)]^n.$$

These are generally nonlinear maps—and since multisymplecticity is a statement about first variations of solutions, let us assume also that these maps are at least C^1 . When $V \times \Sigma$ is infinite-dimensional, we may interpret this as a variational derivative (either the Gâteaux or Fréchet derivative, which are equivalent for C^1 maps, cf. Abraham et al. [2, Corollary 2.10]); in the finite-dimensional case, this is just ordinary continuous differentiability. With ϕ and f defined in this way, it follows that $\phi_\mu^i d\sigma_i^\mu + f_i du^i$ is a C^1 differential 1-form on $V \times \Sigma$, and we say that this 1-form is closed if its exterior derivative vanishes as a 2-form on $V \times \Sigma$.

Definition 5 The flux formulation (11) is *multisymplectic* if solutions satisfy

$$\int_{\partial K} (d\widehat{u}^i \wedge d\widehat{\sigma}_i^\mu) d^{m-1}x_\mu = 0, \quad (12)$$

for all $K \in \mathcal{T}_h$, whenever the C^1 differential 1-form $\phi_\mu^i d\sigma_i^\mu + f_i du^i$ is closed.

Remark 5 The condition (12) is essentially the integral form of the multisymplectic conservation law from Sect. 2.3, where the approximate traces $\widehat{u}$ and $\widehat{\sigma}$ are used instead of the actual traces of u and σ . Note that (12) only needs to hold for $K \in \mathcal{T}_h$, not for arbitrary subdomains $K \subseteq U$ as in (5).

We say that (12) is a *local* multisymplecticity condition because it is a statement only about the local solvers (11a)–(11b) and numerical flux condition (11c) for each $K \in \mathcal{T}_h$. We reserve the *global* question—whether the multisymplectic conservation law holds for arbitrary unions of elements of $\mathcal{T}_h$ —for the next section, where the conservativity condition (11d) will also come into play.

To characterize the multisymplecticity of the flux formulation, we first prove a useful lemma, which relates the boundary integral in (12) to the “jumps” $\widehat{u} - u$ and $\widehat{\sigma} - \sigma$ between the approximate and actual traces on ∂K .

Lemma 2 If $\phi_\mu^i d\sigma_i^\mu + f_i du^i$ is closed and $(u, \sigma, \widehat{u}, \widehat{\sigma}) \in V \times \Sigma \times \widehat{V} \times \widehat{\Sigma}$ satisfies (11a)–(11b) for $K \in \mathcal{T}_h$, then

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

Consequently, the local multisymplecticity condition (12) holds if and only if

$$\int_{\partial K} [\widehat{d}(\widehat{u}^i - u^i) \wedge d(\widehat{\sigma}_i^\mu - \sigma_i^\mu)] d^{m-1}x_\mu = 0. \quad (13)$$

Proof Since (11a) and (11b) hold for all $v \in V(K)$ and $\tau \in \Sigma(K)$, we have

$$\begin{aligned} \int_{\partial K} (\widehat{u}^i d\sigma_i^\mu) d^{m-1}x_\mu &= \int_K [u^i d(\partial_\mu \sigma_i^\mu) + \phi_\mu^i d\sigma_i^\mu] d^m x, \\ \int_{\partial K} (\widehat{\sigma}_i^\mu du^i) d^{m-1}x_\mu &= \int_K [\sigma_i^\mu d(\partial_\mu u^i) - f_i du^i] d^m x, \end{aligned}$$

so taking exterior derivatives gives

$$\begin{aligned} \int_{\partial K} (d\widehat{u}^i \wedge d\sigma_i^\mu) d^{m-1}x_\mu &= \int_K [du^i \wedge d(\partial_\mu \sigma_i^\mu) + d\phi_\mu^i \wedge d\sigma_i^\mu] d^m x, \\ \int_{\partial K} (d\widehat{\sigma}_i^\mu \wedge du^i) d^{m-1}x_\mu &= \int_K [d\sigma_i^\mu \wedge d(\partial_\mu u^i) - df_i \wedge du^i] d^m x. \end{aligned}$$

Subtracting the second equation from the first, the terms involving ϕ and f vanish by the closedness assumption, so we are left with

$$\begin{aligned} \int_{\partial K} (d\widehat{u}^i \wedge d\sigma_i^\mu + du^i \wedge d\widehat{\sigma}_i^\mu) d^{m-1}x_\mu &= \int_K [du^i \wedge d(\partial_\mu \sigma_i^\mu) + d(\partial_\mu u^i) \wedge d\sigma_i^\mu] d^m x \\ &= \int_K \partial_\mu (du^i \wedge d\sigma_i^\mu) d^m x \\ &= \int_{\partial K} (du^i \wedge d\sigma_i^\mu) d^{m-1}x_\mu, \end{aligned}$$

that is,

$$\int_{\partial K} (d\widehat{u}^i \wedge d\sigma_i^\mu + du^i \wedge d\widehat{\sigma}_i^\mu - du^i \wedge d\sigma_i^\mu) d^{m-1}x_\mu = 0.$$

Using this identity, we finally calculate

$$\begin{aligned} \int_{\partial K} [\widehat{d}(\widehat{u}^i - u^i) \wedge d(\widehat{\sigma}_i^\mu - \sigma_i^\mu)] d^{m-1}x_\mu &= \int_{\partial K} (d\widehat{u}^i \wedge d\widehat{\sigma}_i^\mu - d\widehat{u}^i \wedge d\sigma_i^\mu - du^i \wedge d\widehat{\sigma}_i^\mu + du^i \wedge d\sigma_i^\mu) d^{m-1}x_\mu \\ &= \int_{\partial K} (d\widehat{u}^i \wedge d\widehat{\sigma}_i^\mu) d^{m-1}x_\mu, \end{aligned}$$

which completes the proof. $\square$

Equation (13) says that the multisymplecticity of the flux formulation depends entirely on the relationship among u , σ , $\widehat{u}$, and $\widehat{\sigma}$ on ∂K for $K \in \mathcal{T}_h$. That is, it depends entirely on the choice of local flux functions Φ_K .

Definition 6 A local flux function Φ_K is *multisymplectic* if (13) holds whenever $(u, \sigma, \widehat{u}, \widehat{\sigma}) \in V \times \Sigma \times \widehat{V} \times \widehat{\Sigma}$ satisfies (11c).

We now prove multisymplecticity for two particular choices of Φ_K . The first is used for the hybridized Raviart–Thomas (RT-H), Brezzi–Douglas–Marini (BDM-H), and local discontinuous Galerkin (LDG-H) methods; the second is used for the hybridized continuous Galerkin (CG-H) and nonconforming (NC-H) methods. These methods will be discussed further in Sect. 4.

Theorem 1 Suppose that, for all $v \in V(K)$ and $\widehat{v} \in \widehat{V}$, there exists $\widehat{\tau} \in \widehat{\Sigma}(\partial K)$ such that $\widehat{\tau}_i^\mu \mathbf{n}_\mu = \delta_{ij}(\widehat{v}^j - v^j)|_{\partial K}$ for all $i = 1, \dots, n$. Then, for any $\lambda \in L^\infty(\partial K)$ (which is called a “penalty function”), the local flux function

$$\Phi_K(u, \sigma, \widehat{u}, \widehat{\sigma}) = (\widehat{\sigma} - \sigma) - \lambda(\widehat{u} - u)\mathbf{n}$$

is multisymplectic.

Proof The flux condition (11c) says that

$$\int_{\partial K} \delta^{ij}(\widehat{\sigma}_i^\mu - \sigma_i^\mu) \mathbf{n}_\mu \widehat{\tau}_j^\nu d^{m-1}x_\nu = \int_{\partial K} \lambda(\widehat{u}^i - u^i) \widehat{\tau}_i^\mu d^{m-1}x_\mu,$$

for all $\widehat{\tau} \in \widehat{\Sigma}(\partial K)$. By assumption, for any $v \in V(K)$ and $\widehat{v} \in \widehat{V}$, we can choose $\widehat{\tau} \in \widehat{\Sigma}(\partial K)$ such that $\widehat{\tau}_i^\mu \mathbf{n}_\mu = \delta_{ij}(\widehat{v}^j - v^j)|_{\partial K}$ for all $i = 1, \dots, n$, and therefore

$$\int_{\partial K} (\widehat{\sigma}_i^\mu - \sigma_i^\mu)(\widehat{v}^i - v^i) d^{m-1}x_\mu = \int_{\partial K} \lambda \delta_{ij}(\widehat{u}^i - u^i)(\widehat{v}^j - v^j) \mathbf{n}^\mu d^{m-1}x_\mu.$$

Since v and $\widehat{v}$ are arbitrary, this can be written as

$$\int_{\partial K} [(\widehat{\sigma}_i^\mu - \sigma_i^\mu) d(\widehat{u}^i - u^i)] d^{m-1}x_\mu = \int_{\partial K} \lambda \delta_{ij}[(\widehat{u}^i - u^i) d(\widehat{u}^j - u^j)] \mathbf{n}^\mu d^{m-1}x_\mu,$$

and taking the exterior derivative of both sides yields

$$\begin{aligned} & \int_{\partial K} [d(\widehat{\sigma}_i^\mu - \sigma_i^\mu) \wedge d(\widehat{u}^i - u^i)] d^{m-1}x_\mu \\ &= \int_{\partial K} \lambda \delta_{ij} [d(\widehat{u}^i - u^i) \wedge d(\widehat{u}^j - u^j)] \mathbf{n}^\mu d^{m-1}x_\mu. \end{aligned}$$

However, this vanishes by the symmetry of δ and the antisymmetry of $\wedge$, so the multisymplecticity condition (13) holds, as claimed. $\square$

Remark 6 More generally, we can replace $\lambda \delta_{ij} \in L^\infty(\partial K)$ above with penalty functions $\lambda_{ij} \in L^\infty(\partial K)$ such that $\lambda_{ij} = \lambda_{ji}$ for $i, j = 1, \dots, n$, and the argument above still holds. This same generalization applies to the penalty-based HDG methods we will encounter in Sect. 4.

Theorem 2 Suppose that, for all $\tau \in \Sigma(K)$, there exists $\widehat{\tau} \in \widehat{\Sigma}(\partial K)$ such that $\widehat{\tau}_i^\mu \mathbf{n}_\mu = \tau_i^\mu \mathbf{n}_\mu|_{\partial K}$ for $i = 1, \dots, n$. Then the local flux function

$$\Phi_K(u, \sigma, \widehat{u}, \widehat{\sigma}) = (\widehat{u} - u)\mathbf{n}$$

is multisymplectic.

Proof The flux condition (11c) says that

$$\int_{\partial K} (\widehat{u}^i - u^i) \widehat{\tau}_i^\mu d^{m-1}x_\mu = 0,$$

for all $\widehat{\tau} \in \widehat{\Sigma}(\partial K)$. From the assumption on normal traces of elements of $\Sigma(K)$, it follows that

$$\int_{\partial K} (\widehat{u}^i - u^i) (\widehat{\tau}_i^\mu - \tau_i^\mu) d^{m-1}x_\mu = 0,$$

for any $\tau \in \Sigma(K)$ and $\widehat{\tau} \in \widehat{\Sigma}(\partial K)$. This can be written as

$$\int_{\partial K} [(\widehat{u}^i - u^i) d(\widehat{\sigma}_i^\mu - \sigma_i^\mu)] d^{m-1}x_\mu = 0,$$

and taking the exterior derivative yields (13). $\square$

3.3 Global Multisymplecticity Criteria

Whenever (11) is multisymplectic, we may of course sum (12) over an arbitrary collection of elements $\mathcal{K} \subset \mathcal{T}_h$ to obtain the global statement

$$\sum_{K \in \mathcal{K}} \int_{\partial K} (d\widehat{u}^i \wedge d\widehat{\sigma}_i^\mu) d^{m-1}x_\mu = 0, \quad (14)$$

or equivalently, by Lemma 2

$$\sum_{K \in \mathcal{K}} \int_{\partial K} [d(\widehat{u}^i - u^i) \wedge d(\widehat{\sigma}_i^\mu - \sigma_i^\mu)] d^{m-1}x_\mu = 0. \quad (15)$$

In the special case $\mathcal{K} = \mathcal{T}_h$, for a second-order linear elliptic PDE, (15) is precisely Equation 2.10 from Cockburn et al. [13], which establishes the symmetry of the bilinear form used to solve for $\widehat{u}$ once internal degrees of freedom have been eliminated (i.e., the Schur complement). Hence, (15) can be seen as a generalization of this symmetry condition to nonlinear multisymplectic systems and arbitrary $\mathcal{K} \subset \mathcal{T}_h$.

However, (14) is a rather weak global multisymplecticity condition, since it follows trivially from the local condition (12). We now define a stronger version of multisymplecticity, which is more analogous to the classical multisymplectic conservation law (5).

Definition 7 The flux formulation (11) is *strongly multisymplectic* if solutions satisfy

$$\int_{\partial(\bigcup \mathcal{K})} (d\hat{u}^i \wedge d\hat{\sigma}_i^\mu) d^{m-1}x_\mu = 0, \quad (16)$$

for all $\mathcal{K} \subset \mathcal{T}_h$, whenever the C^1 differential 1-form $\phi_\mu^i d\sigma_i^\mu + f_i du^i$ is closed.

Taking $\mathcal{K} = \{K\}$ immediately implies the local condition (12) for each $K \in \mathcal{T}_h$, so the stronger condition (16) indeed implies the weaker condition (14). It follows that (16) holds if and only if the terms of (14) cancel on internal facets. The property that conservation laws “add up” correctly over unions of elements is directly related to the conservativity condition (11d); indeed, this is the reason the term “conservative” is used to describe numerical fluxes with single-valued normal components. (See Equation 3.2 in Arnold et al. [4].)

Let $e = \partial K^+ \cap \partial K^-$ be an internal facet, where $K^\pm \in \mathcal{T}_h$ are distinct. Recall from Remark 3 that an element $\hat{\tau} \in \hat{\Sigma}$ is generally double-valued on e , since the $\hat{\Sigma}(\partial K^+)$ and $\hat{\Sigma}(\partial K^-)$ components need not agree. As is common in the discontinuous Galerkin literature (including Arnold et al. [4] and Cockburn et al. [13]), we define the “normal jump” $[\![\hat{\tau}]\!]_e \in [L^2(e)]^n$ by

$$[\![\hat{\tau}]\!]_e = \hat{\tau}_i^\mu \mathbf{n}_\mu|_{e^+} + \hat{\tau}_i^\mu \mathbf{n}_\mu|_{e^-}.$$

Here, $e^\pm$ denotes that e is oriented according to $\partial K^\pm$, and we take the corresponding component of $\hat{\tau}$ and outer normal $\mathbf{n}$ for each term on the right-hand side. Denoting the set of internal facets of $\mathcal{T}_h$ by $\mathcal{E}_h^\circ := \mathcal{E}_h \setminus \partial U$, we may sum directly over $e \in \mathcal{E}_h^\circ$ to define $[\![\hat{\tau}]\!] \in [L^2(\mathcal{E}_h^\circ)]^n$. Hence, the normal component of $\hat{\tau}$ is single-valued on internal facets if and only if $[\![\hat{\tau}]\!] = 0$.

With this notation, the conservativity condition (11d) can be rewritten as

$$\int_{\mathcal{E}_h^\circ} [\![\hat{\sigma}]\!]_i \hat{v}^i d^{m-1}x = 0, \quad \forall \hat{v} \in \hat{V}_0.$$

If the extension by zero of $[\![\hat{\sigma}]\!]$ to $\mathcal{E}_h$ is in $\hat{V}_0$, then applying the conservativity condition with $\hat{v} = [\![\hat{\sigma}]\!]$ immediately implies $[\![\hat{\sigma}]\!] = 0$, and we say that $\hat{\sigma}$ is *strongly conservative*. However, this is not always the case: for example, if $\hat{\Sigma}(\partial K)$ contains discontinuous traces but $\hat{V}(\partial K)$ contains only continuous traces, then in general $[\![\hat{\sigma}]\!] \notin \hat{V}_0$, so we cannot conclude that $[\![\hat{\sigma}]\!]$ vanishes. In this case, we say that $\hat{\sigma}$ is only *weakly conservative*. (This terminology is taken from Cockburn et al. [13].)

Theorem 3 If $\hat{\sigma} \in \hat{\Sigma}$ satisfies the strong conservativity condition $[\![\hat{\sigma}]\!] = 0$, then for any $\hat{u} \in \hat{V}$ and $\mathcal{K} \subset \mathcal{T}_h$,

$$\sum_{K \in \mathcal{K}} \int_K (d\hat{u}^i \wedge d\hat{\sigma}_i^\mu) d^{m-1}x_\mu = \int_{\partial(\bigcup \mathcal{K})} (d\hat{u}^i \wedge d\hat{\sigma}_i^\mu) d^{m-1}x_\mu.$$

Consequently, if (11) is multisymplectic and strongly conservative, then it is strongly multisymplectic.

Proof Let $e = \partial K^+ \cap \partial K^-$ be an internal facet, where $K^\pm \subset \mathcal{K}$. Since $[\![\widehat{\sigma}]\!] = 0$, we have

$$\int_{e^+} (\widehat{\sigma}_i^\mu d\widehat{u}^i) d^{m-1}x_\mu + \int_{e^-} (\widehat{\sigma}_i^\mu d\widehat{u}^i) d^{m-1}x_\mu = \int_e ([\![\widehat{\sigma}]\!]_i d\widehat{u}^i) d^{m-1}x = 0.$$

Finally, taking the exterior derivative implies that

$$\int_{e^+} (d\widehat{u}^i \wedge d\widehat{\sigma}_i^\mu) d^{m-1}x_\mu + \int_{e^-} (d\widehat{u}^i \wedge d\widehat{\sigma}_i^\mu) d^{m-1}x_\mu = 0,$$

so the contributions from internal facets vanish, as claimed. $\square$

3.4 The Flux Formulation for Exact Solutions

We now apply the theory of the preceding sections to exact solutions of (1), in the sense of distributions, in which case the flux formulation (11) consists of infinite-dimensional function spaces.

Definition 8 The *exact flux formulation* on $\mathcal{T}_h$ is the flux formulation (11) associated with the function spaces

$$\begin{aligned} V(K) &= [H^2(K)]^n, & \Sigma(K) &= [H^1(K)]^{mn}, \\ \widehat{V} &= [L^2(\mathcal{E}_h)]^n, & \widehat{\Sigma}(\partial K) &= [L^2(\partial K)]^{mn}, \end{aligned}$$

along with the flux functions $\Phi_K(u, \sigma, \widehat{u}, \widehat{\sigma}) = \widehat{\sigma} - \sigma$, for $K \in \mathcal{T}_h$.

The next theorem uses a domain-decomposition-type argument to relate the exact flux formulation to solutions of (1), in the sense of distributions, defined over certain global function spaces on U .

Theorem 4 The element $(u, \sigma, \widehat{u}, \widehat{\sigma}) \in V \times \Sigma \times \widehat{V} \times \widehat{\Sigma}$ is a solution to the exact flux formulation if and only if

$$(u, \sigma) \in \left(V \cap [H^1(U)]^n \right) \times \left(\Sigma \cap [H(\operatorname{div}; U)]^n \right)$$

is a solution to (1) on U , in the sense of distributions, and $(\widehat{u}, \widehat{\sigma})$ are the exact traces $\widehat{u} = u|_{\mathcal{E}_h}$ and $\widehat{\sigma}|_{\partial K} = \sigma|_{\partial K}$ for all $K \in \mathcal{T}_h$.

Proof Suppose $(u, \sigma, \widehat{u}, \widehat{\sigma}) \in V \times \Sigma \times \widehat{V} \times \widehat{\Sigma}$ is a solution to the exact flux formulation.

In particular, (11a) holds for all $\tau \in [C_c^\infty(K)]^{mn}$, which immediately gives $\partial_\mu u^i = \phi_\mu^i$ on K , in the sense of distributions. Taking more general test functions $\tau \in [C^\infty(K)]^{mn}$, not necessarily vanishing on ∂K , implies that $\widehat{u}|_{\partial K} = u|_{\partial K}$ in the trace sense. Since this holds for all $K \in \mathcal{T}_h$, it follows that $\widehat{u} = u|_{\mathcal{E}_h}$. Hence, the trace of u is single-valued on $\mathcal{E}_h$, and we may therefore conclude (cf. Brezzi and Fortin [7, Proposition III.1.1]) that $u \in [H^1(U)]^n$.

Similarly, taking smooth test functions v in (11b) implies that $-\partial_\mu \sigma_i^\mu = f_i$ holds in the sense of distributions, and $\widehat{\sigma}_i^\mu \mathbf{n}_\mu|_{\partial K} = \sigma_i^\mu \mathbf{n}_\mu|_{\partial K}$ holds in the trace sense, for all $K \in \mathcal{T}_h$. The flux condition (11c) implies further that $\widehat{\sigma}|_{\partial K} = \sigma|_{\partial K}$, and the conservativity condition (11d) gives $[\![\widehat{\sigma}]\!] = 0$ on $\mathcal{E}_h^\circ$. Hence, the normal trace of σ is single-valued on $\mathcal{E}_h$, and we may therefore conclude (cf. Brezzi and Fortin [7, Proposition III.1.2]) that $\sigma \in [H(\operatorname{div}; U)]^n$.

The converse is a simple verification of (11). If $u \in V \cap [H^1(U)]^n$, then it has a (single-valued) trace $\widehat{u} \in [L^2(\mathcal{E}_h)]^n = \widehat{V}$. Likewise, if $\sigma \in \Sigma \cap [H(\operatorname{div}; U)]^n$, then it has a trace $\widehat{\sigma}|_{\partial K} \in [L^2(\partial K)]^{mn} = \widehat{\Sigma}(\partial K)$ for each $K \in \mathcal{T}_h$; this satisfies $[\![\widehat{\sigma}]\!] = 0$ on $\mathcal{E}_h^\circ$, so (11c)–(11d) hold. Finally, Eqs. (11a)–(11b) hold by the assumption that (u, σ) satisfies (1) in the sense of distributions. $\square$

Corollary 3 *The exact flux formulation satisfies*

$$\int_{\partial(\bigcup \mathcal{K})} (\mathrm{d}u^i \wedge \mathrm{d}\sigma_i^\mu) \mathrm{d}^{m-1}x_\mu = 0,$$

for all $\mathcal{K} \subset \mathcal{T}_h$, whenever $\phi_\mu^i \mathrm{d}\sigma_i^\mu + f_i \mathrm{d}u^i$ is closed.

Proof It follows immediately from Theorem 1 (with $\lambda \equiv 0$) and Theorem 3 that the exact flux formulation is strongly multisymplectic, so it satisfies (16). Moreover, Theorem 4 gives $\widehat{u} = u|_{\mathcal{E}_h}$ and $\widehat{\sigma}|_{\partial K} = \sigma|_{\partial K}$ for all $K \in \mathcal{T}_h$, so we may “remove the hats” from (16). $\square$

4 Multisymplecticity of Particular HDG Methods

We now apply the results of Sect. 3 to the particular HDG methods discussed in Cockburn et al. [13]. Although the flux formulation (11) is more general than that for the class of second-order linear elliptic PDEs they consider, the spaces and fluxes used to define the methods are essentially unchanged.

Throughout this section, we assume that $U \subset \mathbb{R}^m$ is polyhedral and that $\mathcal{T}_h$ is a triangulation of U by m -simplices. We denote by $\mathcal{P}_r(K)$ the space of degree- r polynomials on $K \in \mathcal{T}_h$ and by $\mathcal{P}_r(e)$ the space of degree- r polynomials on $e \in \mathcal{E}_h$. We also define spaces of discontinuous polynomial boundary traces,

$$\mathcal{P}_r(\mathcal{E}_h) := \{\widehat{w} \in L^2(\mathcal{E}_h) : \widehat{w}|_e \in \mathcal{P}_r(e), \forall e \in \mathcal{E}_h^\circ\}$$

and continuous polynomial boundary traces,

$$\mathcal{P}_r^c(\mathcal{E}_h) := \{\widehat{w} \in C^0(\mathcal{E}_h) : \widehat{w}|_e \in \mathcal{P}_r(e), \forall e \in \mathcal{E}_h\},$$

which in Cockburn et al. [13] are called $\mathcal{M}_{h,r}$ and $\mathcal{M}_{h,r}^c$, respectively.

4.1 The RT-H Method

The hybridized Raviart–Thomas (RT-H) method uses the local function spaces

$$V(K) = [\mathcal{P}_r(K)]^n, \quad \Sigma(K) = [\mathcal{P}_r(K)^m + x \mathcal{P}_r(K)]^n,$$

i.e., degree- r Lagrange finite elements and Raviart–Thomas finite elements, respectively, for each $i = 1, \dots, n$. The trace spaces are taken to be

$$\widehat{V} = [\mathcal{P}_r(\mathcal{E}_h)]^n, \quad \widehat{\Sigma}(\partial K) = [L^2(\partial K)]^{mn},$$

and the local flux functions are

$$\Phi_K(u, \sigma, \widehat{u}, \widehat{\sigma}) = \widehat{\sigma} - \sigma.$$

Note that, although $\widehat{\Sigma}(\partial K)$ is infinite-dimensional, the flux condition (11c) simply states that $\widehat{\sigma}|_{\partial K} = \sigma|_{\partial K}$, so we may eliminate this equation and substitute σ wherever $\widehat{\sigma}$ appears in the remaining equations.

Theorem 5 *The RT-H method is strongly multisymplectic.*

Proof Observe that Φ_K is a special case of the flux in Theorem 1, with $\lambda \equiv 0$. Since $\widehat{\Sigma}(\partial K) = [L^2(\partial K)]^{mn}$, it follows that for any $v \in V(K)$ and $\widehat{v} \in \widehat{V}$, we have $(\widehat{v} - v)\mathbf{n}|_{\partial K} \in \widehat{\Sigma}(\partial K)$. Hence, the hypotheses of Theorem 1 are satisfied, so the method is multisymplectic.

To show strong multisymplecticity, let $e \in \mathcal{E}_h^\circ$ be an arbitrary internal facet, and write $e = \partial K^+ \cap \partial K^-$. Since $\sigma|_{K^\pm} \in \Sigma(K^\pm)$, a standard result on Raviart–Thomas elements (cf. Brezzi and Fortin [7, Proposition III.3.2]) implies that $\sigma_i^\mu \mathbf{n}_\mu|_{e^\pm} \in \mathcal{P}_r(e)$ for $i = 1, \dots, n$.¹ The flux condition (11c) implies $\widehat{\sigma}|_{e^\pm} = \sigma|_{e^\pm}$, so it follows from the above that $[\widehat{\sigma}]|_e \in [\mathcal{P}_r(e)]^n$. Since this holds for all $e \in \mathcal{E}_h^\circ$, the extension by zero of $[\widehat{\sigma}]$ to $\mathcal{E}_h$ is in $\widehat{V}_0$. Therefore, the RT-H method is strongly conservative, so Theorem 3 implies that it is strongly multisymplectic. $\square$

Remark 7 To prove multisymplecticity, instead of using Theorem 1, we could simply have used $(\widehat{\sigma} - \sigma)|_{\partial K} = 0$ to see that (13) holds. However, the argument we have developed here is more general, and we will see in Sect. 4.3 that it also applies to methods where $\lambda \neq 0$.

4.2 The BDM-H Method

The hybridized Brezzi–Douglas–Marini (BDM-H) method uses the local function spaces

$$V(K) = [\mathcal{P}_{r-1}(K)]^n, \quad \Sigma(K) = [\mathcal{P}_r(K)]^{mn},$$

¹ This result holds since e lies in an affine hyperplane in $\mathbb{R}^m$, so $x \cdot \mathbf{n}$ is constant on e . It follows that the degree- $(r+1)$ elements of the Raviart–Thomas space nevertheless have degree- r normal traces.

i.e., degree- $(r - 1)$ Lagrange finite elements and degree- r Brezzi–Douglas–Marini finite elements, respectively, for each $i = 1, \dots, n$. As in the RT-H method, the trace spaces are taken to be

$$\widehat{V} = [\mathcal{P}_r(\mathcal{E}_h)]^n, \quad \widehat{\Sigma}(\partial K) = [L^2(\partial K)]^{mn},$$

and the local flux functions are

$$\Phi_K(u, \sigma, \widehat{u}, \widehat{\sigma}) = \widehat{\sigma} - \sigma.$$

As with RT-H, the flux condition (11c) states that $\widehat{\sigma}|_{\partial K} = \sigma|_{\partial K}$, so we may eliminate (11c) and substitute σ for $\widehat{\sigma}$ in the remaining equations.

Theorem 6 *The BDM-H method is strongly multisymplectic.*

Proof Since the trace spaces and flux function are the same as the RT-H method, multisymplecticity follows exactly as in Theorem 5.

Given any internal facet $e = \partial K^+ \cap \partial K^-$, we have $\sigma|_{K^\pm} \in [\mathcal{P}_r(K^\pm)]^{mn}$, so $\widehat{\sigma}|_{e^\pm} = \sigma|_{e^\pm} \in [\mathcal{P}_r(e)]^{mn}$ and therefore $[\widehat{\sigma}]|_e \in [\mathcal{P}_r(e)]^n$. Since this holds for all $e \in \mathcal{E}_h^\circ$, the extension by zero of $[\widehat{\sigma}]$ to $\mathcal{E}_h$ is in $\widehat{V}_0$. Therefore, the BDM-H method is strongly conservative, so Theorem 3 implies that it is strongly multisymplectic. $\square$

4.3 The LDG-H Methods

There are three variants of the hybridized local discontinuous Galerkin method (LDG-H) discussed in Cockburn et al. [13], corresponding to different choices of the local function spaces. In the setting and notation considered here, these three pairs of spaces are:

$$V(K) = [\mathcal{P}_{r-1}(K)]^n, \quad \Sigma(K) = [\mathcal{P}_r(K)]^{mn}, \quad (17a)$$

$$V(K) = [\mathcal{P}_r(K)]^n, \quad \Sigma(K) = [\mathcal{P}_r(K)]^{mn}, \quad (17b)$$

$$V(K) = [\mathcal{P}_r(K)]^n, \quad \Sigma(K) = [\mathcal{P}_{r-1}(K)]^{mn}. \quad (17c)$$

Whichever of these we choose, the trace spaces are

$$\widehat{V} = [\mathcal{P}_r(\mathcal{E}_h)]^n, \quad \widehat{\Sigma}(\partial K) = [L^2(\partial K)]^{mn},$$

and the local flux functions are

$$\Phi_K(u, \sigma, \widehat{u}, \widehat{\sigma}) = (\widehat{\sigma} - \sigma) - \lambda(\widehat{u} - u)\mathbf{n},$$

where the penalty function λ is piecewise constant on ∂K , i.e., constant on each facet. Note that since λ may be different for each $K \in \mathcal{T}_h$, on an internal facet

$e = \partial K^+ \cap \partial K^-$, the constants $\lambda|_{e^\pm}$ need not be equal to one another. The flux condition (11c) states that

$$\widehat{\sigma}|_{\partial K} = [\sigma + \lambda(\widehat{u} - u)\mathbf{n}]|_{\partial K},$$

so we may eliminate (11c) and substitute the expression on the right-hand side wherever $\widehat{\sigma}$ appears in the remaining equations.

Theorem 7 *The LDG-H method is strongly multisymplectic for each of the choices (17a)–(17c) of local function spaces.*

Proof The flux Φ_K is precisely that of Theorem 1. As shown in the proof of Theorem 5, the space $\widehat{\Sigma}(\partial K) = [L^2(\partial K)]^{mn}$ satisfies the hypotheses of Theorem 1, so the LDG-H method is multisymplectic.

Now, for each of (17a)–(17c), we have

$$V(K) \subset [\mathcal{P}_r(K)]^n, \quad \Sigma(K) \subset [\mathcal{P}_r(K)]^{mn}.$$

For any internal facet $e = \partial K^+ \cap \partial K^-$, since $\lambda|_{e^\pm}$ are constants, it follows that

$$\widehat{\sigma}|_{e^\pm} = [\sigma + \lambda(\widehat{u} - u)\mathbf{n}]|_{e^\pm} \in [\mathcal{P}_r(e)]^{mn},$$

so $[\widehat{\sigma}]|_e \in [\mathcal{P}_r(e)]^n$. (Note that if $\lambda|_{e^\pm}$ were arbitrary L^∞ penalty functions, rather than constants, this would not necessarily be true; in fact, $\widehat{\sigma}|_{e^\pm}$ might not be polynomials at all.) Since this holds for all $e \in \mathcal{E}_h^\circ$, the extension by zero of $[\widehat{\sigma}]$ to $\mathcal{E}_h$ is in $\widehat{V}_0$. Therefore, the LDG-H method is strongly conservative, so Theorem 3 implies that it is strongly multisymplectic. $\square$

4.4 The CG-H Method

The hybridized continuous Galerkin (CG-H) method uses the local function spaces

$$V(K) = [\mathcal{P}_r(K)]^n, \quad \Sigma(K) = [\mathcal{P}_{r-1}(K)]^{mn},$$

i.e., degree- r Lagrange finite elements for u^i, v^i and degree- $(r-1)$ Lagrange finite elements for σ_i^μ, τ_i^μ , for $\mu = 1, \dots, m$ and $i = 1, \dots, n$. The trace spaces are taken to be

$$\widehat{V} = [\mathcal{P}_r^c(\mathcal{E}_h)]^n, \quad \widehat{\Sigma}(\partial K) = \{v\mathbf{n}|_{\partial K} : v \in V(K)\},$$

and the local flux functions are

$$\Phi_K(u, \sigma, \widehat{u}, \widehat{\sigma}) = (\widehat{u} - u)\mathbf{n}.$$

Since $u \in V(K)$, we immediately have $u\mathbf{n}|_{\partial K} \in \widehat{\Sigma}(\partial K)$. Moreover, since $\widehat{V}$ consists of continuous polynomials, the degrees of freedom for $\widehat{V}(\partial K)$ are a subset of those for $V(K)$, so $\widehat{u}\mathbf{n}|_{\partial K} \in \widehat{\Sigma}(\partial K)$ as well. Hence, taking the flux condition (11c) with $\widehat{\tau} = (\widehat{u} - u)\mathbf{n}|_{\partial K}$ implies $\widehat{u}|_{\partial K} = u|_{\partial K}$ for all $K \in \mathcal{T}_h$.

Remark 8 The CG-H method is so named because it coincides with the classical continuous Galerkin method with Lagrange finite elements when applied to second-order linear elliptic PDEs of the type considered in Examples 1, 2, and 3, as long as $a = a^{\mu\nu}(x)$ is constant on each $K \in \mathcal{T}_h$. See Cockburn et al. [13,14], Cockburn [12], where the relationship between CG-H and the technique of “static condensation” is also discussed.

More generally, this correspondence also holds for the semilinear system

$$\operatorname{grad} u = a^{-1} \sigma, \quad -\operatorname{div} \sigma = f(\cdot, u).$$

Indeed, since (11c) implies $\widehat{u}|_{\partial K} = u|_{\partial K}$, substituting u for $\widehat{u}$ in (11a) implies $\sigma = a \operatorname{grad} u$, as long as a is constant on each $K \in \mathcal{T}_h$. (Otherwise, $a \operatorname{grad} u$ is generally not in Σ .) It follows that $u \in C^0(U) \cap V$, so for all test functions v in this same space vanishing on ∂U , summing (11b) over $K \in \mathcal{T}_h$ gives

$$\begin{aligned} \int_U a \operatorname{grad} u \cdot \operatorname{grad} v &= \sum_{K \in \mathcal{T}_h} \int_K a \operatorname{grad} u \cdot \operatorname{grad} v \\ &= \sum_{K \in \mathcal{T}_h} \left[\int_K f(\cdot, u) v + \int_{\partial K} \widehat{\sigma} v \cdot \mathbf{n} \right] \\ &= \int_U f(\cdot, u) v + \sum_{K \in \mathcal{T}_h} \int_{\partial K} \widehat{\sigma} v \cdot \mathbf{n} \\ &= \int_U f(\cdot, u) v. \end{aligned}$$

The boundary term vanishes by (11d), since the assumption that v is continuous and vanishes on ∂U implies $v|_{\mathcal{E}_h} \in \widehat{V}_0$. Hence, u is a solution to the continuous Galerkin method for $-\operatorname{div}(a \operatorname{grad} u) = f(\cdot, u)$.

We now prove that the CG-H method is multisymplectic, although—unlike the other HDG methods considered here—it is *not strongly multisymplectic* except in dimension $m = 1$, when multisymplecticity is just ordinary symplecticity.

Theorem 8 *The CG-H method is multisymplectic. It is strongly multisymplectic (i.e., symplectic) when $m = 1$.*

Proof Since $(\widehat{u} - u)|_{\partial K} = 0$, we see directly that (13) holds, so multisymplecticity follows by Lemma 2.

When $m = 1$, facets are simply vertices, so $\mathcal{E}_h$ is discrete and finite, and the continuity conditions on $\widehat{V}$ are trivial. Hence, $[\mathcal{P}_r^c(\mathcal{E}_h)]^n = [\mathcal{P}_r(\mathcal{E}_h)]^n = \mathbb{R}^{|\mathcal{E}_h|n}$. The result follows immediately from the trivial observation that $[\widehat{\sigma}]|_e \in \mathbb{R}^n$ at each internal vertex e . $\square$

Remark 9 Although Φ_K is the same flux function as in Theorem 2, the hypotheses of that theorem do not hold for the CG-H method. Here, $\widehat{\Sigma}(\partial K)$ consists only of $\widehat{\tau}$ whose normal traces are continuous on ∂K , while this is not necessarily true of $\tau|_{\partial K}$ for arbitrary $\tau \in \Sigma(K)$.

Example 5 (CG-H for Laplace’s equation in $\mathbb{R}^2$) Let $m = 2$, $n = 1$, and consider the mixed form of Laplace’s equation,

$$\operatorname{grad} u = \sigma, \quad -\operatorname{div} \sigma = 0.$$

Let us apply the lowest-order CG-H method, with $r = 1$, so that for each $K \in \mathcal{T}_h$, we have $V(K) = \mathcal{P}_1(K)$ and $\Sigma(K) = [\mathcal{P}_0(K)]^2$.

As discussed in Remark 8, we have $\widehat{u}|_{\partial K} = u|_{\partial K}$ and $\sigma|_K = \operatorname{grad} u|_K$. Therefore, $\widehat{\sigma}|_{\partial K}$ is determined by (11b), which in this case states

$$\int_{\partial K} \widehat{\sigma} v \cdot \mathbf{n} = \int_K \operatorname{grad} u \cdot \operatorname{grad} v, \quad \forall v \in V(K).$$

Since $\widehat{\sigma}|_{\partial K} = w\mathbf{n}|_{\partial K}$ for some $w \in V(K)$, we may rewrite this as

$$\int_{\partial K} wv = \int_K \operatorname{grad} u \cdot \operatorname{grad} v, \quad \forall v \in V(K). \quad (18)$$

However, $w|_{\partial K}$ is generally *not* equal to $\operatorname{grad} u \cdot \mathbf{n}|_{\partial K}$, since the latter is piecewise constant (and generally discontinuous) on ∂K , whereas the former must be continuous and linear. Instead, we must set up a linear system and solve for w in terms of u .

For simplicity, let us suppose that K is isometric to the standard, equilateral reference triangle in $\mathbb{R}^3$, defined by

$$T := \{(x, y, z) \in \mathbb{R}_{\geq 0}^3 : x + y + z = 1\}.$$

On T , the Lagrange basis of linear “hat functions” simply consists of the coordinate functions x , y , and z , and we can write

$$u = u_1x + u_2y + u_3z, \quad w = w_1x + w_2y + w_3z.$$

Solving the linear system corresponding to (18) yields, after a calculation,

$$\begin{aligned} w_1 &= \frac{\sqrt{6}}{6}(2u_1 - u_2 - u_3), \\ w_2 &= \frac{\sqrt{6}}{6}(-u_1 + 2u_2 - u_3), \\ w_3 &= \frac{\sqrt{6}}{6}(-u_1 - u_2 + 2u_3). \end{aligned}$$

The multisymplectic form restricted to ∂T is therefore $(du \wedge dw)\mathbf{n}|_{\partial T}$, so the multisymplectic conservation law states that $\int_{\partial T} du \wedge dw = 0$.

Let e_{ij} denote the edge in ∂T going from the i th standard basis vector to the j th standard basis vector in $\mathbb{R}^3$. Using the above expressions for w in terms of u , another calculation shows that

$$\int_{e_{12}} du \wedge dw = \frac{\sqrt{3}}{6} (du_3 \wedge du_1 - du_2 \wedge du_3), \quad (19a)$$

$$\int_{e_{23}} du \wedge dw = \frac{\sqrt{3}}{6} (du_1 \wedge du_2 - du_3 \wedge du_1), \quad (19b)$$

$$\int_{e_{31}} du \wedge dw = \frac{\sqrt{3}}{6} (du_2 \wedge du_3 - du_1 \wedge du_2), \quad (19c)$$

from which it is immediately apparent that

$$\int_{\partial T} du \wedge dw = \int_{e_{12}} du \wedge dw + \int_{e_{23}} du \wedge dw + \int_{e_{31}} du \wedge dw = 0,$$

so the method is indeed multisymplectic on T , following Theorem 8.

We now show that strong multisymplecticity does not hold for the CG-H method when $m > 1$, so the result of Theorem 8 is the best we can hope for. The proof uses a counterexample based on Example 5.

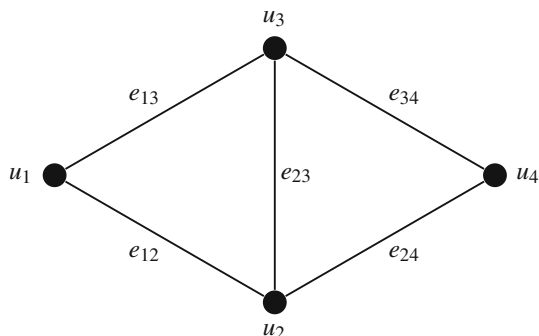
Proposition 1 *The CG-H method is not strongly multisymplectic.*

Proof Consider the mixed form of Laplace’s equation on the domain $U \subset \mathbb{R}^2$ triangulated by two equilateral triangles, as shown in Fig. 1. For simplicity, as in Example 5, we suppose that each of these is isometric to the standard reference triangle T , so that the edge lengths are all $\sqrt{2}$. Note that $\hat{V}_0 = \{0\}$, since the degrees of freedom for $\hat{V}$ all lie on ∂U , so the conservativity condition (11d) is trivial. Hence, there are no constraints on $(u, \sigma, \hat{u}, \hat{\sigma})$ other than the local conditions (11a)–(11c) on each triangle, as discussed in Example 5.

From the calculation in (19), we have

$$\begin{aligned} \int_{e_{12}} du \wedge dw + \int_{e_{13}} du \wedge dw &= \frac{\sqrt{3}}{6} (du_3 \wedge du_1 - du_1 \wedge du_2) \\ &= \frac{\sqrt{3}}{6} (du_2 + du_3) \wedge du_1, \end{aligned}$$

Fig. 1 A triangulation on which the CG-H method is not strongly multisymplectic for Laplace’s equation



and similarly,

$$\int_{e_{34}} du \wedge dw + \int_{e_{24}} du \wedge dw = \frac{\sqrt{3}}{6} (du_2 + du_3) \wedge du_4,$$

so adding these gives

$$\int_{\partial U} du \wedge dw = \frac{\sqrt{3}}{6} (du_2 + du_3) \wedge (du_1 + du_4) \neq 0. \quad (20)$$

We conclude that this is nonzero since $u_1, \dots, u_4$ are independent degrees of freedom that may be varied independently. In particular, if we take variations $(v, \text{grad } v, v, \widehat{\tau})$ and $(v', \text{grad } v', v', \widehat{\tau}')$ with

$$v_2 = v'_1 = 1, \quad v_1 = v_3 = v_4 = v'_2 = v'_3 = v'_4 = 0,$$

we have

$$\int_{\partial U} (v\widehat{\tau}' - v'\widehat{\tau}) \cdot \mathbf{n} = \frac{\sqrt{3}}{6} [(v_2 + v_3)(v'_1 + v'_4) - (v'_2 + v'_3)(v_1 + v_4)] = \frac{\sqrt{3}}{6} \neq 0.$$

Hence, the CG-H method is not strongly multisymplectic. $\square$

Remark 10 The failure of strong multisymplecticity, in this example, may also be seen via the failure of the bilinear form $(v, v') \mapsto \int_{\partial U} v\widehat{\tau}' \cdot \mathbf{n}$ to be symmetric. Indeed, if we write $\mathbf{v} = (v_1, v_2, v_3, v_4)$ and $\mathbf{v}' = (v'_1, v'_2, v'_3, v'_4)$, then we may represent this as the quadratic form,

$$\int_{\partial U} v\widehat{\tau}' \cdot \mathbf{n} = \frac{\sqrt{3}}{6} \mathbf{v}^T \begin{bmatrix} 2 & -1 & -1 & 0 \\ 0 & 2 & -2 & 0 \\ 0 & -2 & 2 & 0 \\ 0 & -1 & -1 & 2 \end{bmatrix} \mathbf{v}',$$

which is immediately seen not to be symmetric. Neglecting the scalar factor of $\sqrt{3}/6$, the antisymmetrization of this matrix is

$$\begin{aligned} \begin{bmatrix} 2 & -1 & -1 & 0 \\ 0 & 2 & -2 & 0 \\ 0 & -2 & 2 & 0 \\ 0 & -1 & -1 & 2 \end{bmatrix} - \begin{bmatrix} 2 & -1 & -1 & 0 \\ 0 & 2 & -2 & 0 \\ 0 & -2 & 2 & 0 \\ 0 & -1 & -1 & 2 \end{bmatrix}^T &= \begin{bmatrix} 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & -1 & -1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}. \end{aligned}$$

This is precisely the matrix corresponding to

$$(du_2 + du_3) \otimes (du_1 + du_4) - (du_1 + du_4) \otimes (du_2 + du_3) = (du_2 + du_3) \wedge (du_1 + du_4),$$

which agrees with (20).

Remark 11 Cockburn et al. [13] observe that CG-H can be seen as a limiting case of the LDG-H method (17c) where the penalty $\lambda \equiv +\infty$. They also consider more general LDG-H methods where λ is infinite on some elements and finite on others [13, Section 3.4]. However, allowing the penalty to be infinite imposes continuity conditions on $\widehat{V}$, meaning that conservativity (and thus multisymplecticity) is only weak rather than strong.

4.5 The NC-H Method

The hybridized nonconforming (NC-H) method uses the local function spaces

$$V(K) = [\mathcal{P}_r(K)]^n, \quad \Sigma(K) = [\mathcal{P}_{r-1}(K)]^{mn},$$

just as in the CG-H method. The trace spaces are

$$\widehat{V} = [\mathcal{P}_{r-1}(\mathcal{E}_h)]^n, \quad \widehat{\Sigma}(\partial K) = \{\widehat{w}\mathbf{n} : \widehat{w}|_e \in [\mathcal{P}_{r-1}(e)]^n, \forall e \in \partial K\},$$

and the local flux functions are

$$\Phi_K(u, \sigma, \widehat{u}, \widehat{\sigma}) = (\widehat{u} - u)\mathbf{n}.$$

Unlike CG-H, $(\widehat{u} - u)\mathbf{n}|_{\partial K}$ is generally not in $\widehat{\Sigma}(\partial K)$, so we cannot conclude that $\widehat{u}$ equals u on ∂K , except in the weak sense of (11c).

Theorem 9 *The NC-H method is strongly multisymplectic.*

Proof The flux Φ_K is that considered in Theorem 2, so it suffices to show that the hypothesis of that theorem holds, i.e.: for all $\tau \in \Sigma(K)$, there exists $\widehat{\tau} \in \widehat{\Sigma}(\partial K)$ such that $\widehat{\tau}_i^\mu \mathbf{n}_\mu = \tau_i^\mu \mathbf{n}_\mu|_{\partial K}$ for $i = 1, \dots, n$. Given $\tau \in \Sigma(K)$, this condition is satisfied by $\widehat{\tau}_i^\mu = \tau_i^\nu \mathbf{n}_\nu \mathbf{n}^\mu|_{\partial K}$, i.e., the projection of $\tau|_{\partial K}$ onto the unit normal, so the NC-H method is multisymplectic.

Next, if $e = \partial K^+ \cap \partial K^-$ is an interior facet, then the definition of $\widehat{\Sigma}(\partial K^\pm)$ implies that $\widehat{\sigma}|_{e^\pm} \in [\mathcal{P}_{r-1}(e)]^n$, so $[\widehat{\sigma}]|_e \in [\mathcal{P}_{r-1}(e)]^n$. Since this holds for all $e \in \mathcal{E}_h^\circ$, the extension by zero of $[\widehat{\sigma}]$ to $\mathcal{E}_h$ is in $\widehat{V}_0$. Therefore, the NC-H method is strongly conservative, so Theorem 3 implies that it is strongly multisymplectic. $\square$

4.6 The IP-H Methods

We finally consider the special case of the hybridized interior penalty method (IP-H), which—unlike the methods considered above—is somewhat idiosyncratic to semilinear elliptic systems. Consider a system of the form

$$\partial_\mu u^i = a_{\mu\nu}^{ij} \sigma_j^\nu, \quad -\partial_\mu \sigma_i^\mu = \frac{\partial F}{\partial u^i}, \quad (21)$$

which generalizes the scalar ($n = 1$) semilinear PDEs discussed in Sect. 2 to $n \geq 1$. Here, $a = a_{ij}^{\mu\nu}(x)$ is a symmetric, positive-definite $mn \times mn$ matrix with inverse $a_{\mu\nu}^{ij}(x) := (a_{ij}^{\mu\nu}(x))^{-1}$. These are the de Donder–Weyl equations for the Hamiltonian

$$H(x, u, \sigma) = \frac{1}{2} a_{\mu\nu}^{ij}(x) \sigma_i^\mu \sigma_j^\nu + F(x, u),$$

so in particular (21) is a canonical multisymplectic system of PDEs.

For such a system, the IP-H method uses the local function spaces

$$V(K) = [\mathcal{P}_r(K)]^n, \quad \Sigma(K) = [\mathcal{P}_r(K)]^{mn}. \quad (22a)$$

We also consider the “IP-H-like” method, suggested by Cockburn et al. [13, p. 1351], which uses the function spaces

$$V(K) = [\mathcal{P}_r(K)]^n, \quad \Sigma(K) = [\mathcal{P}_{r-1}(K)]^{mn}. \quad (22b)$$

For both methods, the trace spaces are taken to be

$$\widehat{V} = [\mathcal{P}_r(\mathcal{E}_h)]^n, \quad \widehat{\Sigma}(\partial K) = [L^2(\partial K)]^{mn}.$$

Based on the observation that the classical solution satisfies $\sigma_i^\mu = (a \operatorname{grad} u)_i^\mu = a_{ij}^{\mu\nu} \partial_\nu u^j$, these methods take the local flux functions to be

$$\Phi_K(u, \sigma, \widehat{u}, \widehat{\sigma}) = (\widehat{\sigma} - a \operatorname{grad} u) - \lambda(\widehat{u} - u)\mathbf{n}.$$

Thus, the IP-H and “IP-H-like” methods are essentially the LDG-H methods (17b) and (17c), respectively, except with σ replaced by $a \operatorname{grad} u$ in the local flux functions. As with LDG-H methods, the penalty function $\lambda|_{\partial K}$ is taken to be piecewise constant on ∂K for each $K \in \mathcal{T}_h$, and on internal facets $e = \partial K^+ \cap \partial K^-$, the constants $\lambda|_{e^\pm}$ need not be equal to one another. The flux condition (11c) states that

$$\widehat{\sigma}_i^\mu|_{\partial K} = [a_{ij}^{\mu\nu} \partial_\nu u^j + \lambda \delta_{ij} (\widehat{u}^j - u^j) \mathbf{n}^\mu]|_{\partial K},$$

so we may eliminate (11c) and substitute the expression on the right-hand side wherever $\widehat{\sigma}$ appears in the remaining equations. We also assume, following Cockburn et al. [13], that a is constant on each $K \in \mathcal{T}_h$.

Theorem 10 *For the semilinear system (21) where a is constant on each $K \in \mathcal{T}_h$, the IP-H and “IP-H-like” methods are strongly multisymplectic.*

Proof To prove that the methods are multisymplectic, Lemma 2 states that it suffices to show that (13) is satisfied on each $K \in \mathcal{T}_h$. Observe that

$$(\widehat{\sigma}_i^\mu - \sigma_i^\mu)|_{\partial K} = [(a_{ij}^{\mu\nu} \partial_\nu u^j - \sigma_i^\mu) + \lambda \delta_{ij} (\widehat{u}^j - u^j) \mathbf{n}^\mu]|_{\partial K}.$$

We have previously seen that

$$\int_{\partial K} \lambda \delta_{ij} [d(\widehat{u}^i - u^i) \wedge d(\widehat{u}^j - u^j)] \mathbf{n}^\mu d^{m-1} x_\mu = 0,$$

by the symmetry of δ and the antisymmetry of $\wedge$. Therefore, it remains to show that the terms involving $a_{ij}^{\mu\nu} \partial_\nu u^j - \sigma_i^\mu$ vanish as well. Integrating (11a) by parts gives

$$\begin{aligned} \int_{\partial K} (\widehat{u}^i - u^i) \tau_i^\mu d^{m-1} x_\mu &= \int_K (\phi_\mu^i - \partial_\mu u^i) \tau_i^\mu d^m x \\ &= \int_K (a_{\mu\nu}^{ij} \sigma_j^\nu - \partial_\mu u^i) \tau_i^\mu d^m x, \end{aligned}$$

and since this holds for all $\tau \in \Sigma(K)$, we can write

$$\int_{\partial K} [(\widehat{u}^i - u^i) d\sigma_i^\mu] d^{m-1} x_\mu = \int_K [(a_{\mu\nu}^{ij} \sigma_j^\nu - \partial_\mu u^i) d\sigma_i^\mu] d^m x.$$

Taking the exterior derivative of both sides yields

$$\begin{aligned} \int_{\partial K} [d(\widehat{u}^i - u^i) \wedge d\sigma_i^\mu] d^{m-1} x_\mu &= \int_K [a_{\mu\nu}^{ij} d\sigma_j^\nu \wedge d\sigma_i^\mu - d(\partial_\mu u^i) \wedge d\sigma_i^\mu] d^m x \\ &= \int_K [-d(\partial_\mu u^i) \wedge d\sigma_i^\mu] d^m x, \end{aligned}$$

where $a_{\mu\nu}^{ij} d\sigma_j^\nu \wedge d\sigma_i^\mu = 0$ by the symmetry of a and the antisymmetry of $\wedge$. On the other hand, taking $\tau = a \operatorname{grad} v$ for $v \in V(K)$, we have

$$\begin{aligned} \int_{\partial K} (\widehat{u}^i - u^i) \tau_i^\mu d^{m-1} x_\mu &= \int_K (a_{\mu\nu}^{ij} \sigma_j^\nu - \partial_\mu u^i) a_{ik}^{\mu\lambda} \partial_\lambda v^k d^m x \\ &= \int_K (\sigma_i^\mu - a_{ij}^{\mu\nu} \partial_\nu u^j) \partial_\mu v^i d^m x, \end{aligned}$$

and since this holds for all $v \in V(K)$, we can write

$$\int_{\partial K} [(\widehat{u}^i - u^i) d(a_{ij}^{\mu\nu} \partial_\nu u^j)] d^{m-1} x_\mu = \int_K [(\sigma_i^\mu - a_{ij}^{\mu\nu} \partial_\nu u^j) d(\partial_\mu u^i)] d^m x.$$

Note that we need a to be constant on K in order for $\tau = a \operatorname{grad} v \in [\mathcal{P}_{r-1}(K)]^{mn} \subset \Sigma(K)$ to be an admissible test function for both methods (22a)–(22b). Taking the exterior derivative of both sides yields

$$\begin{aligned}
& \int_{\partial K} [\widehat{u}^i - u^i] \wedge d(a_{ij}^{\mu\nu} \partial_\nu u^j)] d^{m-1} x_\mu \\
&= \int_K [d\sigma_i^\mu \wedge d(\partial_\mu u^i) - a_{ij}^{\mu\nu} d(\partial_\nu u^j) \wedge d(\partial_\mu u^i)] d^m x \\
&= \int_K [d\sigma_i^\mu \wedge d(\partial_\mu u^i)] d^m x \\
&= \int_{\partial K} [\widehat{u}^i - u^i] \wedge d\sigma_i^\mu d^{m-1} x_\mu,
\end{aligned}$$

where again we have used the symmetry of a and the antisymmetry of $\wedge$. Rearranging this last equality yields

$$\int_{\partial K} [\widehat{u}^i - u^i] \wedge d(a_{ij}^{\mu\nu} \partial_\nu u^j - \sigma_i^\mu) d^{m-1} x_\mu = 0,$$

which is precisely (13). Hence, the IP-H method is multisymplectic.

To show strong multisymplecticity, let $e = \partial K^+ \cap \partial K^-$ be an internal facet. Then

$$\widehat{\sigma}|_{e^\pm} = [a \operatorname{grad} u + \lambda(\widehat{u} - u)\mathbf{n}]|_{e^\pm} \in [\mathcal{P}_r(e)]^{mn},$$

so $\llbracket \widehat{\sigma} \rrbracket|_e \in [\mathcal{P}_r(e)]^n$. (Here, we use that $a|_{e^\pm}$ and $\lambda|_{e^\pm}$ are constant.) Hence, $\llbracket \widehat{\sigma} \rrbracket$ may be extended by zero to an element of $\widehat{V}_0$, so both methods are strongly conservative and thus, by Theorem 3, strongly multisymplectic. $\square$

5 Conclusion

We have generalized the flux formulation and HDG framework of Cockburn et al. [13] to the much larger family of canonical systems of PDEs (1), which includes nonlinear systems. Within this framework, we have established the multisymplecticity of several HDG methods, when such methods are applied to canonical Hamiltonian systems (2). These methods include “hybridized” versions of several widely used classes of finite element methods, suggesting that—when multisymplectic structure preservation is desired—these general-purpose methods may be used instead of the specialized methods constructed in previous work, which are often limited in their order of accuracy and/or use on unstructured meshes.

It is perhaps not surprising that so many finite element methods are multisymplectic. Indeed, like the Ritz–Galerkin method, the multisymplectic conservation law is intimately related to variational principles (cf. Lawruk et al. [24]; Kijowski and Tulczyjew [23]; Marsden et al. [27]; Gotay et al. [18,19]; Vankerschaver et al. [42]). However, any construction of multisymplectic finite element methods must address two difficulties: first, that the multisymplectic conservation law holds for smooth solutions, while finite element spaces are non-smooth; and second, that finite element basis functions can have support on several elements, posing an obstacle to expressing a localized, per-element conservation law. The flux formulation of Sect. 3 provides a solution to

both of these difficulties. Hybridization (i.e., introducing separate spaces of boundary traces and fluxes) provides a way to express the multisymplectic conservation law on individual elements, while the exact flux formulation of Sect. 3.4 provides a bridge between the smooth and non-smooth cases (particularly Theorem 4 and Corollary 3).

There are two final points we wish to emphasize about these results, by comparison to the previous work discussed in Sect. 1.

1. Multisymplectic HDG methods may be applied to systems of the form (1), whether or not the user is aware of any canonical multisymplectic/Hamiltonian structure. Yet, if such a structure is present, it will automatically be preserved. This is analogous to the case for certain classes of symplectic integrators for ODEs (e.g., symplectic partitioned Runge–Kutta methods), but contrasts with previously studied multisymplectic methods for PDEs on unstructured meshes (e.g., Lagrangian variational methods).
2. Many of the previously constructed multisymplectic methods merely satisfy a “discrete version” of the multisymplectic conservation law, e.g., a finite-difference version of (4) on a lattice. This is true even for the previous work on multisymplectic finite element methods [11,21,44], which are only “multisymplectic” in a finite-difference sense on the lattice of degrees of freedom. Moreover, this discrete multisymplectic conservation law may differ from method to method, depending on how the divergence operator is discretized.

By contrast, the multisymplectic conservation law (12) satisfied by these HDG methods is *exactly* the integral multisymplectic conservation law (5), restricted to the numerical traces and fluxes on ∂K for each element $K \in \mathcal{T}_h$, while (16) extends this to arbitrary unions of elements. Furthermore, this multisymplectic conservation law does not differ from method to method: it has precisely the same form for weak solutions, in the exact flux formulation of Sect. 3.4, as it does for each of the HDG methods of Sect. 4.

One direction for future work is the application of multisymplectic HDG methods to Hamiltonian time-evolution PDEs. There are two natural ways in which such work might proceed. First, a multisymplectic HDG method might be used to semidiscretize the PDE in space, resulting in a finite-dimensional system of ODEs. Sánchez et al. [41] have recently shown that, when an LDG-H method is used to semidiscretize the acoustic wave equation, the resulting system of ODEs is Hamiltonian, and one may then apply a symplectic integrator in time. Second, one might consider the application of spacetime HDG methods, as in Rhebergen and Cockburn [39,40], Griesmaier and Monk [20]. In this approach, one would simultaneously discretize space and time by applying the flux formulation of Sect. 3 to the case where U is a spacetime domain and $\mathcal{T}_h$ a decomposition into spacetime elements $K \in \mathcal{T}_h$. The results of Sect. 3 are formulated in sufficient generality to include such methods, so one might apply them to investigate the multisymplecticity of specific spacetime HDG methods.

Acknowledgements This research was supported in part by the Marsden Fund of the Royal Society of New Zealand and by the Simons Foundation (Award #279968 to Ari Stern). We also wish to thank the anonymous referees for their helpful comments and suggestions.

References

1. Abraham, R., Marsden, J.E.: Foundations of mechanics. Benjamin/Cummings Publishing Co. Inc. Advanced Book Program, Reading, Mass. (1978). Second edition, revised and enlarged, With the assistance of Tudor Rațiu and Richard Cushman
2. Abraham, R., Marsden, J.E., Ratiu, T.: Manifolds, tensor analysis, and applications. *Applied Mathematical Sciences*, vol. 75. Second edn. Springer-Verlag, New York (1988)
3. Agoshkov, V.I.: Poincaré-Steklov's operators and domain decomposition methods in finite-dimensional spaces. In: First International Symposium on Domain Decomposition Methods for Partial Differential Equations (Paris, 1987), pp. 73–112. SIAM, Philadelphia, PA (1988)
4. Arnold, D.N., Brezzi, F., Cockburn, B., Marini, L.D.: Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.* **39**(5), 1749–1779 (2001/2002)
5. Ascher, U.M., McLachlan, R.I.: Multisymplectic box schemes and the Korteweg–de Vries equation. *Applied Numerical Mathematics* **48**(3–4), 255–269 (2004)
6. Belishev, M., Sharafutdinov, V.: Dirichlet to Neumann operator on differential forms. *Bull. Sci. Math.* **132**(2), 128–145 (2008)
7. Brezzi, F., Fortin, M.: Mixed and hybrid finite element methods. Springer Series in Computational Mathematics, vol. 15. Springer-Verlag, New York (1991)
8. Bridges, T.J.: Symplecticity, reversibility and elliptic operators. In: Nonlinear dynamical systems and chaos (Groningen, 1995), *Progr. Nonlinear Differential Equations Appl.*, vol. 19, pp. 1–20. Birkhäuser, Basel (1996)
9. Bridges, T.J.: Multi-symplectic structures and wave propagation. *Math. Proc. Cambridge Philos. Soc.* **121**(1), 147–190 (1997)
10. Bridges, T.J., Reich, S.: Multi-symplectic integrators: numerical schemes for Hamiltonian PDEs that conserve symplecticity. *Phys. Lett. A* **284**(4–5), 184–193 (2001)
11. Chen, J.B.: Variational integrators and the finite element method. *Appl. Math. Comput.* **196**(2), 941–958 (2008)
12. Cockburn, B.: Static condensation, hybridization, and the devising of the HDG methods. In: Building bridges: connections and challenges in modern approaches to numerical partial differential equations, *Lect. Notes Comput. Sci. Eng.*, vol. 114, pp. 129–177. Springer (2016)
13. Cockburn, B., Gopalakrishnan, J., Lazarov, R.: Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.* **47**(2), 1319–1365 (2009)
14. Cockburn, B., Gopalakrishnan, J., Wang, H.: Locally conservative fluxes for the continuous Galerkin method. *SIAM J. Numer. Anal.* **45**(4), 1742–1776 (2007)
15. de Donder, T.: Théorie Invariantive du Calcul des Variations. Second edn. Gauthier-Villars (1935)
16. Frank, J., Moore, B.E., Reich, S.: Linear PDEs and numerical methods that preserve a multisymplectic conservation law. *SIAM J. Sci. Comput.* **28**(1), 260–277 (2006)
17. Gotay, M.J.: A multisymplectic framework for classical field theory and the calculus of variations. I. Covariant Hamiltonian formalism. In: Mechanics, analysis and geometry: 200 years after Lagrange, North-Holland Delta Ser., pp. 203–235. North-Holland, Amsterdam (1991)
18. Gotay, M.J., Isenberg, J., Marsden, J.E.: Momentum maps and classical relativistic fields. Part II: Canonical analysis of field theories (2004a). [arXiv:math-ph/0411032](https://arxiv.org/abs/math-ph/0411032)
19. Gotay, M.J., Isenberg, J., Marsden, J.E., Montgomery, R.: Momentum maps and classical relativistic fields. Part I: Covariant Field Theory (2004b). [arXiv:physics/9801019v2](https://arxiv.org/abs/physics/9801019v2) [math-ph]
20. Griesmaier, R., Monk, P.: Discretization of the wave equation using continuous elements in time and a hybridizable discontinuous Galerkin method in space. *J. Sci. Comput.* **58**(2), 472–498 (2014)
21. Guo, H.Y., Ji, X.M., Li, Y.Q., Wu, K.: A note on symplectic, multisymplectic scheme in finite element method. *Commun. Theor. Phys. (Beijing)* **36**(3), 259–262 (2001)
22. Hairer, E., Lubich, C., Wanner, G.: Geometric numerical integration, *Springer Series in Computational Mathematics*, vol. 31. Second edn. Springer-Verlag, Berlin (2006)
23. Kijowski, J., Tulczyjew, W.M.: A symplectic framework for field theories, *Lecture Notes in Physics*, vol. 107. Springer-Verlag, Berlin-New York (1979)
24. Lawruk, B., Śniatycki, J., Tulczyjew, W.M.: Special symplectic spaces. *J. Differential Equations* **17**, 477–497 (1975)
25. Lew, A., Marsden, J.E., Ortiz, M., West, M.: Asynchronous variational integrators. *Arch. Ration. Mech. Anal.* **167**(2), 85–146 (2003)

26. Marsden, J.E., Hughes, T.J.R.: Mathematical foundations of elasticity. Dover Publications, Inc., New York (1994). Corrected reprint of the 1983 original
27. Marsden, J.E., Patrick, G.W., Shkoller, S.: Multisymplectic geometry, variational integrators, and non-linear PDEs. *Comm. Math. Phys.* **199**(2), 351–395 (1998)
28. Marsden, J.E., Pekarsky, S., Shkoller, S., West, M.: Variational methods, multisymplectic geometry and continuum mechanics. *J. Geom. Phys.* **38**(3–4), 253–284 (2001)
29. Marsden, J.E., Shkoller, S.: Multisymplectic geometry, covariant Hamiltonians, and water waves. *Math. Proc. Cambridge Philos. Soc.* **125**(3), 553–575 (1999)
30. McDonald, F., McLachlan, R.I., Moore, B.E., Quispel, G.: Travelling wave solutions of multisymplectic discretizations of semi-linear wave equations. *Journal of Difference Equations and Applications* **22**(7), 913–940 (2016)
31. McLachlan, R., Offen, C.: Preservation of bifurcations of Hamiltonian boundary value problems under discretisation (2018a). [arXiv:1804.07468](https://arxiv.org/abs/1804.07468) [math.NA]
32. McLachlan, R.I., Offen, C.: Bifurcation of solutions to Hamiltonian boundary value problems. *Non-linearity* **31**(6), 2895 (2018b)
33. McLachlan, R.I., Ryland, B.N., Sun, Y.: High order multisymplectic Runge–Kutta methods. *SIAM Journal on Scientific Computing* **36**(5), A2199–A2226 (2014)
34. McLachlan, R.I., Sun, Y., Tse, P.: Linear stability of partitioned Runge–Kutta methods. *SIAM Journal on Numerical Analysis* **49**(1), 232–263 (2011)
35. Moore, B., Reich, S.: Backward error analysis for multi-symplectic integration methods. *Numer. Math.* **95**(4), 625–652 (2003)
36. Partridge, P.W., Brebbia, C.A., Wrobel, L.C.: The dual reciprocity boundary element method. *International Series on Computational Engineering. Computational Mechanics Publications*, Southampton; copublished with Elsevier Applied Science, London (1992)
37. Reich, S.: Finite volume methods for multi-symplectic PDEs. *BIT* **40**(3), 559–582 (2000a)
38. Reich, S.: Multi-symplectic Runge-Kutta collocation methods for Hamiltonian wave equations. *J. Comput. Phys.* **157**(2), 473–499 (2000b)
39. Rhebergen, S., Cockburn, B.: A space-time hybridizable discontinuous Galerkin method for incompressible flows on deforming domains. *J. Comput. Phys.* **231**(11), 4185–4204 (2012)
40. Rhebergen, S., Cockburn, B.: Space-time hybridizable discontinuous Galerkin method for the advection-diffusion equation on moving and deforming meshes. In: *The Courant-Friedrichs-Lewy (CFL) condition*, pp. 45–63. Birkhäuser/Springer, New York (2013)
41. Sánchez, M.A., Ciucu, C., Nguyen, N.C., Peraire, J., Cockburn, B.: Symplectic Hamiltonian HDG methods for wave propagation phenomena. *J. Comput. Phys.* **350**, 951–973 (2017)
42. Vankerschaver, J., Liao, C., Leok, M.: Generating functionals and Lagrangian partial differential equations. *J. Math. Phys.* **54**(8), 082,901, 22 (2013)
43. Weyl, H.: Geodesic fields in the calculus of variation for multiple integrals. *Ann. of Math. (2)* **36**(3), 607–629 (1935)
44. Zhen, L., Bai, Y., Li, Q., Wu, K.: Symplectic and multisymplectic schemes with the simple finite element method. *Phys. Lett. A* **314**(5–6), 443–455 (2003)