SPATIAL DISCRETIZATION OF PARTIAL DIFFERENTIAL EQUATIONS WITH INTEGRALS

ROBERT I. MCLACHLAN

Abstract. We consider the problem of constructing spatial finite difference approximations on a fixed, arbitrary grid, which have analogues of any number of integrals of the partial differential equation and of some of its symmetries. A basis for the space of of such difference operators is constructed; most cases of interest involve a single such basis element. (The "Arakawa" Jacobian is such an element.) We show how the topology of the grid affects the complexity of the operators.

1. Conservative discretization

"Numerical methods for nonlinear conservation laws are among the great success stories of modern numerical analysis." (Iserles [3])

Conservative discretizations of partial differential equations have been explored for a long time. What does "conservative" mean? An early definition is due to Lax and Wendroff [9], who considered the class on PDEs with one spatial dimension,

$$u_t + \partial_x(f(u)) = 0 \tag{1.1}$$

and called discretizations of the form

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} = \frac{H(u_{i+j}^n, \ldots, u_{i-j+1}^n) - H(u_{i+j-1}^n, \ldots, u_{i-j}^n)}{\Delta x} \tag{1.2}$$

conservative. See [3] for an introduction to such methods. More generally, the formulation (1.1) is called conservative, and the expanded form

$$u_t + f'(u)u_x$$

nonconservative, with these terms carrying over to the corresponding discrete forms. The PDE (1.1) reflects, amongst other things, conservation of the integral of u (e.g. total mass, momentum, etc.), and the conservative discretization (1.2) preserves a discrete analog: $\sum_i u_i^{n+1} = \sum_i u^n$. The full consequences for the discrete scheme of the form (1.2) remain unclear.

More recently the term has been applied to PDEs that can be written purely in terms of intrinsic differential operators such as div, grad, and curl. A conservative spatial discretization is then one which preserves discrete analogues of these operators' integral identities (e.g., Stokes's theorem). In many cases these obey maximum principles and have robust stability properties in difficult situations such as rough grids and discontinuous coefficients [20].

Research at MSRI is supported in part by NSF grant DMS-9701755.

Schemes have also been developed for particular equations that inherit conserved quantities approximating those of the PDE. An early and famous example is the Arakawa Jacobian [1], a discretization of $v_x w_y - v_y w_x$ which, when applied to the two-dimensional Euler fluid equations, provides two conservation laws corresponding to energy and enstrophy, both quadratic functions. It is widely used in computational meteorology. There are many energy-conserving schemes for particular PDEs: Fei and Vásquez [4] for the sine-Gordon equation; Glassey [6] for the Zakharov equations; Glassey and Schaeffer [7] for a nonlinear wave equation. The original presentations of all these are somewhat ad-hoc, the proof of conservation relying on a telescoping sum.

The Arakawa Jacobian has the extremely nice property that it can be applied to systems (in two space dimensions, with two variables) with *any* two integrals, not just energy and enstrophy. It was further explained and generalized to arbitrary grids by Salmon and Talley in [19]. It is this systematic approach that we generalize in this paper to equations with any number of integrals, space dimensions, and variables. Our formulation includes all integral-preserving discretizations.

Having integrals of course reduces the evolution to a smaller space, which, when their level sets are compact, gives the method a form or nonlinear stability. Often, more is true: Preservation of a discrete form of $\int u\,dx$ by the Lax-Wendroff form (1.2) leads to correct shock speeds, and preservation of energy and enstrophy by the Arakawa Jacobian prevents energy cascading to small length scales [1].

The Euler equations, the sine-Gordon equation and so on are examples of *Hamiltonian* PDEs, which suggests that one should look for semi- or fully-discrete forms which preserve not only a discrete energy but also a discrete Hamiltonian (symplectic) structure. For systems with canonical Hamiltonian structure, this possibility was explored in [11]. But even before the importance of Hamiltonian PDEs was widely recognized, for which a watershed event was perhaps the 1983 conference [22], it had been pointed out by Morrison [16] that spatial discretizations of non-canonical Hamiltonian PDEs will not normally be Hamiltonian. One example apart, the curious 'sine bracket' Hamiltonian discretization of the Jacobian [24], this is a difficult and essentially unsolved problem. Probably the right generalization of 'Hamiltonian' has not yet been found.

We are thus reluctantly led to consider only energy-conserving discretizations. Or perhaps we should not be reluctant: Simo et al. [21] have argued and presented detailed evidence from elastodynamics that conserving energy leads to excellent nonlinear stability properties that preserving symplectic structure does not. (Essentially because symplectic schemes can only brake the fast modes, whereas energy-conserving schemes can also damp them.)

In Hamiltonian systems, energy is normally seen as playing a distinguished role. Yet there may be other conserved quantities just as important for the long-time dynamics. Some of them, the 'Casimir' integrals, can be due to the Hamiltonian structure itself. Non-Hamiltonian systems can also have conserved quantities. Even in the ODE example of the free rigid body, the relationship between schemes preserving energy and/or momentum and/or symplectic structure is quite complicated [10].

In this paper we go some way towards uniting these different integrals and different points of view. Our goal is to develop a methodology for building spatial discretizations that preserve discrete analogues of any given set of integrals. It should be systematic, all-inclusive, and reproduce known schemes. We do this in a formulation in which the integrals appear explicitly; the integrals themselves can then be discretized in any way. The basic "finite difference molecule" is now a completely skew-symmetric tensor. Symmetry plays a fundamental role, and we will see how the skew-symmetry of this tensor interacts with other desired symmetries of the scheme (e.g., translational and rotational invariance) in a nontrivial way.

2. Hamiltonian and other PDEs with integrals

We consider PDEs with independent spatial variables $x \in \mathbb{R}^d$ and dependent variables $u(x) \in \mathbb{R}^m$. We loosely call m the "number of variables." The relevant class of sufficiently smooth real-valued functionals of $u(x)$ will be denoted $\mathcal{U}$. A Hamiltonian PDE is specified by a Hamiltonian $\mathcal{H} \in \mathcal{U}$ and a Poisson bracket $\{\,,\,\} : \mathcal{U} \times \mathcal{U} \to \mathcal{U}$:

$$\dot{u} = \{u, \mathcal{H}\}, \tag{2.3}$$

where the Poisson bracket is bilinear, skew-symmetric

$$\{\mathcal{F}, \mathcal{G}\} = -\{\mathcal{G}, \mathcal{F}\}, \tag{2.4}$$

and obeys the Jacobi identity

$$\{\mathcal{F}, \{\mathcal{G}, \mathcal{H}\}\} + \{\mathcal{G}, \{\mathcal{H}, \mathcal{F}\}\} + \{\mathcal{H}, \{\mathcal{F}, \mathcal{G}\}\} = 0 \tag{2.5}$$

and the Leibniz rule

$$\{\mathcal{F}, \mathcal{G}\mathcal{H}\} = \{\mathcal{F}, \mathcal{G}\}\mathcal{H} + \{\mathcal{F}, \mathcal{H}\}\mathcal{G} \tag{2.6}$$

for all $\mathcal{F}$, $\mathcal{G}$, $\mathcal{H} \in \mathcal{U}$. In fact, these axioms imply the existence of a Hamiltonian (or 'Poisson') operator $\mathcal{J}$, such that the Poisson bracket takes the form

$$\{\mathcal{F}, \mathcal{G}\} = \int \frac{\delta \mathcal{F}}{\delta u} \mathcal{J} \frac{\delta \mathcal{G}}{\delta u}\, dx \tag{2.7}$$

and (2.3) becomes

$$\dot{u} = \mathcal{J} \frac{\delta \mathcal{H}}{\delta u} \tag{2.8}$$

A particularly important example is the conservation law (1.1), which takes this form with x, $u \in R$ and

$$\mathcal{H} = \int F(u(x))\, dx, \quad \mathcal{J} = \partial_x, \quad F' = f. \tag{2.9}$$

The system (2.3) has functional $\mathcal{I} \in \mathcal{U}$ as an integral if $\dot{\mathcal{I}} = \{\mathcal{I}, \mathcal{H}\}$=0. Some integrals $\mathcal{C}$ are distinguished in that $\{C, \mathcal{F}\} = 0$ for all $\mathcal{F} \in \mathcal{U}$; they are called Casimirs. The operator $\mathcal{J} = \partial_x$ has a single Casimir, $\mathcal{C} = \int u\, dx$, because $\mathcal{J}(\delta\mathcal{C}/\delta u) = \partial_x 1 = 0$.

The peculiarly Hamiltonian character of these PDEs is due to the Jacobi identity satisfied by the Poisson bracket. If, as argued previously, we discard this identity, what class of systems result? Energy is still conserved, because of the skew-symmetry of the bracket. Systems (2.3) may still have integrals and the operator $\mathcal{J}$ may still have Casimirs. Do such systems retain any other special properties? The answer is no: *all* systems with an integral $\mathcal{H}$ can be written in the form (2.3) for some choice of the bracket (or equivalently, for some choice of the skew-adjoint operator $\mathcal{J}$). So to study systems with an integral, and their discretizations, we may without loss of generality assume the form (2.3).

This is most easily seen in the finite dimensional case, as shown recently by Quispel and Capel [18]. Take $u \in \mathbb{R}^n$ as coordinates on phase space. Poisson ($\equiv$ non-canonical Hamiltonian) systems

$$\dot{u} = \{u, H\} = J(u)\nabla H(u)$$

have integral H. But suppose an arbitrary system $\dot{u} = f(u)$ has integral H. Let $z = \nabla H$ and

$$J_{ij} = \frac{f_i z_j - f_j z_i}{\sum_k z_k^2} \tag{2.10}$$

Then J is skew-symmetric, and $J\nabla H = f$ as required. This J is singular at critical points of H, but in [13] it is shown that if f and H are smooth, and the critical points of H are nondegenerate, then there is a smooth matrix J such that $f = J\nabla H$.

This idea extends easily to systems with any number of integrals: The system of ODEs $\dot{u} = f(x)$ has integrals $I^1, \dots, I^p$ if and only if there exists a totally skew-symmetric $(p+1)$-tensor K such that for all x where the vectors ∇I^i are linearly independent,

$$f_i = K_{ijk\dots} \frac{\partial I^1}{\partial x_j} \frac{\partial I^2}{\partial x_k} \dots. \tag{2.11}$$

For, suppose K exists. Then $\dot{I}^j = f \cdot \nabla I^j = 0$, so each I^j is an integral. Conversely, suppose f has integrals I^j. Then (using exterior algebra, see [2])

$$K = \frac{f \wedge \nabla I^1 \cdots \wedge \nabla I^p}{\det(\nabla I^i \cdot \nabla I^j)}$$

satisfies (2.11). K is determined uniquely only in the case $n = p + 1$; see [13] for further details. We write the inner product (2.11) as

$$f = K(\nabla I^1, \nabla I^2, \dots).$$

What about Casimirs? Suppose that instead of contracting K against all the integrals, as in (2.11), we contract against just one, say I^1. Then $\widetilde{K} = K(\nabla I^1)$ is a skew p-tensor which has I^1 as a Casimir, in the sense that $\widetilde{K}(\nabla I^1) \equiv 0$ and any differential equation formed from this $\widetilde{K}$ (as in (2.11)) will have I^1 as an integral. But there are many different K's, all generating the same system $\dot{u} = f$, that do not have I^1 as a Casimir. Thus the distinction between the Hamiltonian, other integrals, and Casimirs, that was present for Hamiltonian systems, is lost now. We are free to move between different representations of f as needed.

(There is another importance difference. If $J \in \mathbb{R}^{n\times n}$ satisfies the Jacobi identity and has locally constant rank m, then J automatically has $n-m$ Casimirs [17]. This need not be true if the Jacobi identity does not hold: there may not be $n-m$ functions whose gradients span J's nullspace. This is another reason for constructing J's as above that automatically have the required Casimirs.)

One can use the tensor K to define a $(p+1)$-bracket,

$$\{F_1,\ldots,F_{p+1}\} = K(\nabla F_1,\ldots,\nabla F_{p+1})$$

which is multilinear, a derivation in each argument, and completely antisymmetric. Such brackets have been revived in modern times by Nambu [15], who, amongst other things, introduced the 3-bracket on $\mathbb{R}^3$ given by $K_{ijk} = \varepsilon_{ijk}$ (the alternating tensor). This gives systems of the form

$$\dot{u} = K(\nabla I^1, \nabla I^2) = \nabla I^1 \times \nabla I^2.$$

In particular, the free rigid body takes this form with $I^1 = \frac{1}{2}|u|^2$ being total angular momentum and $I^2 = \frac{1}{2}\sum u_i^2/A_i$ being the kinetic energy, where the A_i are the body's moments of inertia. Contracting against ∇I^1 gives the standard, Lie-Poisson form of the equations,

$$\dot{u}_i = J_{ik}(u)\frac{\partial I^2}{u_k} = \varepsilon_{ijk} u_j \frac{\partial I^2}{\partial u_k}.$$

However, later studies [5, 23], attempting to build a true generalization of Hamiltonian mechanics from such brackets, have found that not all constant K's satisfy the required "fundamental identity" (the analogue of the Jacobi identity); its solutions all lead to systems with $n-1$ integrals. So it is not clear that interesting dynamics as well as interesting algebraic structure will arise in this way.

The situation for PDEs is formally the same. For example, for a PDE $\dot{u} = f(u)$ with one integral $\mathcal{I}$, one can define the operator $\mathcal{J}$ analogously to (2.10) by

$$\mathcal{J}(x,x')v(x') = \frac{\int \left(f(x)z(x') - f(x')z(x)\right) v(x')\,dx'}{\int \frac{\delta \mathcal{I}}{\delta u}(x')z(x')\,dx'}$$

where $z(x)$ is any smooth function, and

$$f(x) = \int \mathcal{J}(x,x')\frac{\delta I}{\delta u}(x')\,dx'.$$

The extension to multiple integrals $\{\mathcal{I}^j\}$ is similar. However, questions of convergence of the integrals arise, and the nonuniqueness situation is much worse: it is not clear how to construct *local* operators, for example. However, as our goal is to construct finite-dimensional finite difference operators, the representation (2.11) is sufficient.

Looking back at Eq. (1.1), we see that it encompasses two important conservation laws expressed in two different ways. The fact that $\partial_x(\delta u/\delta u) = \partial_x 1 = 0$ means that the Casimir $\int u\,dx$ is an integral, and the conservative scheme (1.2) maintains a discrete analog of this. In this example, this property is relatively easy to preserve under discretization: a system has $\sum_i u_i$ as an integral if and only if it can be written in the form $\dot{u}_i = J_{ij}F_j$, where J is not necessarily skew, but $\sum_i J_{ij} = 0$ for all j. Without loss of generality we can take J to be in the form (1.2), with just two nonzero diagonals. The form of the F_j chosen in (1.2) is necessary for translation invariance.

Secondly, if f is a variational derivative, $f = \delta\mathcal{H}/\delta u$ say, then the skew-adjointness of $\mathcal{J} = \partial_x$ means that $\mathcal{H}$ is an integral. To preserve this property under discretization means taking (again, without loss of generality) $\dot{u}_i = J_{ij}\frac{\partial H}{\partial u_j}$, where J is skew symmetric. Note that such a J need not *a priori* have $\sum u_i$ as a Casimir, and, similarly, the nonsymmetric J used in (1.2) does not preserve any discrete H. Thus, the two expressions of conservation laws are in fact independent.

In this paper we generalize the second form. The first form is deceptively simple in this example, because the Casimir is so simple. It is not clear how to modify (1.2) to incorporate different Casimirs. In the second form the integral appears explicitly and, once J is found, *any* quantity can be conserved.

Before continuing, we mention one trivial but complete solution to the whole problem. Why not contract K against *all* the integrals and have simply $\dot{u}_i = f_i$, with all I^j being integrals of f? That is, $f \cdot \nabla I^j = 0$ for all j. This case is already included in the above formulation, with f regarded as a "skew 1-tensor." The equations $f \cdot \nabla I^j = 0$ are linear and can be solved in many ways, for example, by starting with an arbitrary f and projecting to the subspace $\{f \cdot \nabla I^j = 0 \; j = 1, \dots, p\}$. One objection is that this solution is so general that it is not clear how to proceed in any particular case. For example, to modify f as little as possible one might choose orthogonal projection, but this will couples all of the f_i. By incorporating more of the known structure of the problem we can work more systematically.

3. Method of discretization

We wish to construct discretizations of the form (2.11). There are two ways to proceed. One could take a particular PDE, write it in the form $\dot{u} = \mathcal{K}(\delta I^1, \dots)$, and discretize this skew-adjoint operator $\mathcal{K}$, preserving skew symmetry. This is difficult, if only because such formulations of PDEs are new and have not been widely developed yet. Instead, we study systems of the form (2.11) in their own right, constructing elementary tensors $K_{ijk\dots}$, and seeing what PDEs they can be used to approximate. That is, we establish (in a sense defined below)

$$K(v^1, \dots, v^p) \cdots = \mathcal{K}(v^1, \dots, v^p) + \mathcal{O}(h^r).$$

We call K a *finite difference tensor* and $K(v^1, \dots, v^p)$ a *finite difference.* Then, the integrals $\mathcal{I}^i$ can be discretized in any way, say by

$$I^i(u) = \mathcal{I}^i(u) + \mathcal{O}(h^r)$$

(this amounts to a numerical quadrature) giving the conservative system of ODEs

$$\dot{u} = K(\nabla I^1, \dots, \nabla I^p). \tag{3.12}$$

Since this form includes all systems with integrals I^i, we can be confident of not missing any in our construction.

We start with an elementary example illustrating how easy it is to break skew symmetry. With one integral in one space dimension, we are seeking an antisymmetric matrix K. On a constant-spaced grid, central differences have

$$K = \frac{1}{2h}\begin{pmatrix} \ddots & & & & \\ & -1 & 0 & 1 & \\ & & -1 & 0 & 1 \\ & & & & \ddots \end{pmatrix}, \tag{3.13}$$

which is antisymmetric. On a non-constant-spaced grid, if we let $p(x)$ be the quadratic interpolating (x_0, v_0), (x_1, v_1), and (x_2, v_2) and use the estimate $v_1' = p'(x_1)$, the associated matrix K is not antisymmetric—it even has a nonzero diagonal. Taking its antisymmetric part is not a good idea, as we have no idea what operator K^{T} approximates. Indeed, it is not immediately clear what bandwidth is required to achieve order 2, say, with an antisymmetric matrix. The element $K_{i,i-1}$, which is relevant to $\dot{u}_i$, must also contribute to $\dot{u}_{i-1}$.

Nonconstant operators also pose problems. Let $\mathcal{K} = u\partial_x + \partial_x u$. Central differences are not skew, but it is not obvious that the skew matrix

$$K = \frac{1}{h}\begin{pmatrix} \ddots & & & & \\ & -u_{i-1} & 0 & u_i & \\ & & -u_i & 0 & u_{i+1} \\ & & & & \ddots \end{pmatrix}, \tag{3.14}$$

is a discretization of $\mathcal{K}$, or how to increase its order from 1.

Below we develop some requirements on the tensors K, and construct all the elementary ones, for various numbers of integrals, space dimensions, and grids.

4. Definitions & theory

The fundamental objects are the *grid* L, the *index set* M, the *symmetry group* G, and the skew tensors $K \in \Lambda^{p+1}(\mathbb{R}^M)$, which we now define.

Let L be a set of indices of grid points. To each index $i \in L$ there is a physical point $x_i \in \mathbb{R}^d$. Let $\{1, \ldots, m\}$ be the set of indices of the dependent variables, so that the full, discrete state space is indexed by the *index set*

$$M = L \times \{1, \ldots, m\}.$$

A *grid function* is a real function on M; for example, the system state is given by the grid function $u : M \to \mathbb{R}$. Its value at point $i = (l, \alpha) \in M$ is written $u_i = u_{(l,\alpha)}$. (That is, we are assembling all the unknowns into a big "column vector.") When $m = 1$, we drop the second subscript entirely.

For simplicity, we only consider the interpretation of this function in which $u_{(l,\alpha)} \approx u_\alpha(x_l)$. (Staggered grids and u_i representing other functionals of $u(x)$ do not affect our main line of argument.)

The p functions, and their corresponding grid functions, which are to be inserted in (3.12) are denoted $v^1, \ldots, v^p$. However, when $p = 1$ we denote it v, and when $p = 2$ we denote them v and w, to reduce the number of indices.

A discretization of a PDE is thus a vector field on $\mathbb{R}^M$, and a discretization of an operator $\mathcal{K}$ is a skew $(p+1)$-tensor $K_{i_0 \ldots i_p} = K_{\boldsymbol{i}} = K_{(\boldsymbol{l},\boldsymbol{\alpha})} \in \mathbb{R}$, $i_j \in M$; i.e.,

$K \in \Lambda^{p+1}(\mathbb{R}^M)$. Since we are trying to *construct* such tensors, intermediate steps will also involve nonskew tensors, i.e., real functions on M^{p+1}.

K *approximates* $\mathcal{K}$ *to order* r if

$$K_{i_0\ldots i_p} v^1(x_{i_1})\ldots v^p(x_{i_p}) = \mathcal{K}(v_1,\ldots,v^p)(x_{i_0}) + \mathcal{O}(h^r)$$

for all smooth $v^1,\ldots,v^p$. We sometimes drastically abbreviate this to $Kv = \mathcal{K}v + \mathcal{O}(h^r)$. We also abbreviate $v^1_{i_1}\ldots v^p_{i_p}$ to $v_{\boldsymbol{i}}$.

Let the grid L have a nonnegative distance function $|j-k|$. This extends to M by $|(j,\alpha)-(k,\beta)| := |j-k|$. The *bandwidth* of K is the smallest c such that $K_{\boldsymbol{i}} = 0$ for all $\boldsymbol{i}$ such that $|i_j - i_k| > c$.

Two examples are the Euclidean distance $|x_i - x_j|$, giving the "Euclidean bandwidth," and the minimum number of edges traversed going from i and j, where the grid points have been connected to form a graph, giving the "graph bandwidth." For example, K in Eq. (3.14) has graph bandwidth 1. K is *local* if it has a finite bandwidth even on infinite grids.

The grid L may be structured, like a square or triangular lattice, or unstructured. Let G be a symmetry group acting on $M = L \times \{1,\ldots,m\}$. We usually consider only *spatial symmetries*, those which are the identity in their second slot, i.e. $\pi_2 g(l,\alpha) = \alpha$. These merely rearrange grid functions on L.

Furthermore, G is usually a subgroup of the symmetry group of the continuous physical space. For example, suppose this space is the plane. Many PDEs of physical interest are invariant under the group $E(2)$ of Euclidean motions of the plane (the semi-direct product of rotations, reflections, and translations, $O(2) \circledS \mathbb{R}^2$). A discrete version of such a PDE can inherit some of this invariance if L has a subgroup of $E(2)$ as a symmetry group. Examples are the square lattice, which has $D_4 \circledS \mathbb{Z}^2$ as a symmetry group (8 rotations and reflections, plus discrete translations) and the equilateral triangular lattice, which has $D_3 \circledS \mathbb{Z}^2$. For the Euclidean group of the sphere there are no such natural lattices, and the dislocations that occur, e.g., when triangulating an icosahedral grid, are known to destabilize numerical methods and create artifacts in the solutions.

Some nonspatial symmetries can also be included. For example, for a PDE involving $f(v_{2x} - v_{1y})$ we might include the map $(x,y,v_1,v_2) \mapsto (y,x,v_2,v_1)$ in G.

Order of accuracy can also sometimes be expressed as a symmetry. One way to ensure second-order accuracy is for the expansion of the discretization error in powers of the spatial grid size h to have only odd or only even terms present. This is equivalent to being invariant under the operation $h \mapsto -h$, or $x \mapsto -x$. This can only apply if $x \mapsto -x$ is a symmetry of the lattice itself, which it is for square and triangular lattices.

To include this possibility we equip each element of G with a sign, $\mathrm{sgn}(g) = \pm 1$, such that G is homomorphic to $\mathbb{Z}_2$. The map corresponding to $h \mapsto -h$ would then have sign 1 (-1) when the operator has an even (odd) number of derivatives.

The action of G extends to an action on $\Lambda^{p+1}(\mathbb{R}^M)$ by

$$gK_{i_0\ldots i_p} := K_{g(i_0)\ldots g(i_p)}$$

which we write as

$$(gK)_{\boldsymbol{i}} = g(K_{\boldsymbol{i}}) = K_{g(\boldsymbol{i})}.$$

A tensor K is *G-invariant* if $gK = (\operatorname{sgn} g)K$ for all $g \in G$.

Thus we have the following requirements on the finite difference operator K:

- K should be completely skew-symmetric;
- K should be G-invariant;
- K should be as simple as possible;
- K should approximate the desired continuous operator to the desired order;
- K should be local.

However, these requirements conflict with each other.

One way to construct operators such as K with the required symmetry properties is to sum over the symmetry group. Given any tensor K,

$$\sum_{\sigma\in S_{p+1},g\in G} \operatorname{sgn}(\sigma)\operatorname{sgn}(g)K_{\sigma(g(\boldsymbol{i}))} \tag{4.15}$$

is completely skew-symmetric and G-invariant. This suggests two ways to construct symmetric K's:

- Start with a K which approximates the desired continuous operator, and symmetrize it;
- Start with a very simple K, such as a basis element for the space of $(p+1)$-tensors, symmetrize it, and see what continuous operator it approximates.

A major drawback of the first strategy is that we have no control over what the symmetrized K approximates.

The second strategy builds a "library" of all such difference operators, from which linear combinations can be taken as desired. However, the form (4.15) is not convenient for writing down these operators in the usual way, which requires the coefficients of each v_k^j appearing in the resulting grid function at a particular point i_0. That is, we want to know $K_{i_0 i_1 \dots i_p}$ for a particular i_0. We derive finite differences in this form in three stages: firstly, for $m=1$ variable; secondly, for $m \ge 1$ variables with no unknowns at the same grid point coupled; thirdly, the general case, $m \ge 1$ variables with arbitrary coupling.

Case 1. $m = 1$ **variable.** When $m=1$ we drop the second component of the index $i = (l,\alpha)$. We can take $i_0 = 0$. Fix a multi-index $\boldsymbol{i} \in M^{p+1}$ where $i_0 = 0$ and start with the elementary tensor defined by $K_{\boldsymbol{i}} = 1$, $K_{\boldsymbol{j}} = 0$ for all $\boldsymbol{j} \ne \boldsymbol{i}$. We assume that the indices in $\boldsymbol{i}$ are distinct, for otherwise skew-symmetrizing K would lead to the zero tensor. Skew-symmetrizing K gives a tensor of bandwidth $\max_{j,k} |i_j - i_k|$. It is

$$H_{\boldsymbol{l}} = \sum_{\rho\in S_{p+1},h\in G} \operatorname{sgn}(\rho)\operatorname{sgn}(h)K_{h(\rho(\boldsymbol{l}))}$$

so the vector field at the point 0 is $\sum_{\boldsymbol{l}} H_{\boldsymbol{l}} v_{\boldsymbol{l}}$ where $l_0 = 0$. Since K is a discrete delta function, there is only one nonzero term in this sum, i.e.,

$$\sum_{\boldsymbol{l}} H_{\boldsymbol{l}} v_{\boldsymbol{l}} = \sum_{\rho,h,\boldsymbol{l}} \operatorname{sgn}(\rho)\operatorname{sgn}(h) v_{\boldsymbol{l}}, \tag{4.16}$$

where the sum is taken over all ρ, h, and $\boldsymbol{l}$ such that $\rho(h(\boldsymbol{l})) = \boldsymbol{i}$ and $l_0 = 0$. Therefore, $\boldsymbol{l} = \rho^{-1}(h^{-1}(\boldsymbol{i}))$. Let $g = h^{-1}$, $\sigma = \rho^{-1}$ so that Eq. (4.16) becomes

$$\sum_{\sigma,g} \operatorname{sgn}(\sigma)\operatorname{sgn}(g) v_{\sigma(g(\boldsymbol{i}))} \tag{4.17}$$

where the sum is over all σ and g such that $\sigma(g(\boldsymbol{i}))_0 = 0$.

Since no two indices in $\boldsymbol{i}$ are the same, for each such g, let σ be such that $\sigma(g(\boldsymbol{i}))_0 = 0$; then the remaining σ's that satisfy this equation lie in S_p, the permutations of the last p indices. The sum over S_p can be evaluated to give a determinant, giving the vector field at point 0,

$$F(\boldsymbol{i}) := \sum_{g\in G_{\boldsymbol{i}}} \operatorname{sgn}(\sigma)\operatorname{sgn}(g) \det V_{\sigma(g(\boldsymbol{i}))} \tag{4.18}$$

where

$$G_{\boldsymbol{i}} = \{g \in G : \exists \sigma \in S_{p+1} \text{ such that } \sigma(g(\boldsymbol{i}))_0 = 0\}, \tag{4.19}$$

and the $p \times p$ matrix V has (j,k) entry

$$(V_{\boldsymbol{l}})_{jk} = v^j_{l_k}, \quad 1 \le j,k \le p. \tag{4.20}$$

(The first subscript of $l_0 = 0$ and does not appear in the matrix.)

We introduce a graphical notation for formulas such as Eq. (4.18). An arrow connecting gridpoints $i_1, \ldots, i_p$ will indicate a term $\det V_{\boldsymbol{i}}$. For $p > 1$, the sign factors can be incorporated by applying a permutation of sign $\operatorname{sgn}(\sigma)\operatorname{sgn}(g)$ to the i_j (for $p = 2$ and $p = 3$ we merely change the direction of the arrow if the required sign is -1, equivalent to writing the columns of V in reverse order) or by choosing σ in (4.19) so that $\operatorname{sgn}(\sigma)\operatorname{sgn}(g) = 1$. The reader is encouraged to refer immediately to Fig. (1(a)) and its associated finite difference Eq. (5.27) to see how easy this is.

Thus, constructing skew finite differences amounts to choosing an initial arrow, finding the group $G_{\boldsymbol{i}}$, and finding the image of the initial arrow under $G_{\boldsymbol{i}}$.

Case 2. $m \ge 1$ **variables, distinct points coupled.** Let $\boldsymbol{i} = (\boldsymbol{l}, \boldsymbol{\alpha})$ is the chosen basis element, where $l_0 = 0$. "Distinct points coupled" means that $l_j \ne l_k$ for all j, k.

The single element $\boldsymbol{i}$ will contribute to the vector field at $(0, \alpha_j)$ for all j; therefore, we will construct, not a single basis vector field, but the family of vector fields spanned by $K_{(\boldsymbol{l},\boldsymbol{\alpha})}$ *for all* $\boldsymbol{\alpha}$. That is, we allow coupling of all components right from the start. We introduce the rank $p+1$, dimension m tensor $T_{\boldsymbol{\alpha}} \in \mathbb{R}^{m^{p+1}}$ and start with $K_{\boldsymbol{i},\boldsymbol{\alpha}} = T_{\boldsymbol{\alpha}}$, $K_{\boldsymbol{j},\boldsymbol{\alpha}} = 0$ for $\boldsymbol{j} \ne \boldsymbol{i}$.

Passing from Eq. (4.17) to Eq. (4.18) only required that the l_j be distinct; therefore the symmetrized vector field at the point 0 (corresponding to Eq. (4.18) in the single variable case) is

$$\sum_{\substack{g\in G_{\boldsymbol{i}} \\ \alpha_0,\ldots,\alpha_p}} \operatorname{sgn}(\sigma)\operatorname{sgn}(g) T_{\boldsymbol{\alpha}} \det V_{\sigma(g(\boldsymbol{i}))} \frac{\partial}{\partial u_{\sigma(g(\boldsymbol{\alpha}))_0}} \tag{4.21}$$

We want to find the α_0-component of this vector field. To do this we relabel the dummy indices $\boldsymbol{\alpha}$ by applying $(\sigma g)^{-1}$ *in the second slot only* to get the vector field at point 0 in component α_0,

$$F(\boldsymbol{i}) := \sum_{\substack{g\in G_{\boldsymbol{i}}\\ \alpha_1,\ldots,\alpha_p}} \operatorname{sgn}(\sigma)\operatorname{sgn}(g) T_{g^{-1}(\sigma^{-1}(\boldsymbol{\alpha}))} \det V_{(\sigma(g(\boldsymbol{l})),\boldsymbol{\alpha})} \tag{4.22}$$

Notice that in (4.22), each determinant involves the same components of the v^i. Also, if g is a spatial symmetry, then $g^{-1}(\sigma^{-1}(\boldsymbol{\alpha})) = \sigma^{-1}(\boldsymbol{\alpha})$.

The diagram notation extends easily to (4.22). To the arrow $\sigma(g(\boldsymbol{l}))$ we attach the label $g^{-1}(\sigma^{-1}(\boldsymbol{\alpha}))$ indicating the T-tensor attached to that determinant.

Case 3. $m \geq 1$ variables, arbitrary coupling. Equality of some of the l_j affects the sum over permutations in Eq. (4.17). Let $\boldsymbol{j} = h(\rho(\boldsymbol{l}))$ where $j_0 = 0$. Let $n(\boldsymbol{j})$ be the number of 0's in $\boldsymbol{j}$. Then the subgroup of S_{p+1} leaving $\sigma(\boldsymbol{j})_0 = 0$ is not S_p as it was before. It is convenient to have a sum of determinants of all $p \times p$ matrices, so we express this subgroup as the product of S_p and the flips $(0k)$, $k = 1, \ldots, n(\boldsymbol{j}) - 1$. This subgroups of S_{p+1} contains all permutations of the first n elements; summing over these merely skew-symmetrizes T. We could have imposed this in the first place for simplicity. Since there will usually be g_j (a translation, say) with $g_j(l_j) = 0$, this is true for any set of equal elements of $\boldsymbol{l}$.

Let $\Sigma_l = \{\sigma : \sigma(\boldsymbol{l}) = \boldsymbol{l}\}$ be the symmetry group of $\boldsymbol{l}$. To sum up, we

$$\text{take } T \text{ to skew-symmetric under } \Sigma_{\boldsymbol{l}}. \tag{4.23}$$

With this assumption, each flip $(0k)$ gives an equal contribution and we can evaluate the sum over permutations to give the symmetrized vector field at the point 0

$$F(\boldsymbol{i}) := \sum_{\substack{g\in G_{\boldsymbol{i}}\\ \alpha_1,\ldots,\alpha_p}} \operatorname{sgn}(\sigma)\operatorname{sgn}(g) n(g(\boldsymbol{l})) T_{g^{-1}(\sigma^{-1}(\boldsymbol{\alpha}))} \det V_{(\sigma(g(\boldsymbol{l})),\boldsymbol{\alpha})} \tag{4.24}$$

In the diagrams, to the arrow $\sigma(g(\boldsymbol{l}))$ we attach the weight $n(g(\boldsymbol{l}))$.

To summarize, the final finite difference evaluated at the point $(0, \alpha_0)$, is given by Equations (4.24) (4.19), and (4.20). Eq. (4.24) specializes to Eq. (4.22) when all elements of $\boldsymbol{l}$ are distinct, and specializes further to (4.18) when $m = 1$. In practice, from a diagram one writes down the finite difference directly from its diagram.

We develop these diagrams and study the resulting differences for different numbers of integrals and dimensions of phase space, and different symmetry groups G.

5. Examples

Case 1. $p = 1$ integral, $d = 1$ space dimension. By scaling it is sufficient to consider $\boldsymbol{i} = (0, 1)$. The bandwidth is 1. The group $G_{\boldsymbol{i}}$ has the single element $i \mapsto i - 1$. The permutation which brings 0 to the front is $(0, -1) \mapsto (-1, 0)$, with sign -1. Thus we get the standard central difference

$$F(0, 1) = v_1 - v_{-1}.$$

On a grid with constant spacing h,

$$v_1 - v_{-1} = 2hv_x + \mathcal{O}(h^3)$$

.

All other examples are related to this one:

1. by Richardson extrapolation,
$$8F(0,1) - F(0,2) = 12hv_x + \mathcal{O}(h^5).$$
2. eliminating the leading order term(s) gives finite differences approximating higher-order differential operators:
$$F(0,2) - 2F(0,1) = 2h^3 v_{xxx} + \mathcal{O}(h^5).$$
3. taking a linear combination of these basis elements gives stencils that approximate other first-order differential operators. This is equivalent to multiplying by a symmetric tensor s_{ij}. To get a smooth continuous limit we can take, e.g., $s_{ij} = q(x_i, x_j, u_i, u_j)$, where q is symmetric in its first and second pairs of arguments. If K is the tensor corresponding to the vector field $F(0,1)$,
$$s_{ij} K_{ij} v_j = 2h(s\partial_x + \partial_x s)v + \mathcal{O}(h^3), \tag{5.25}$$
where $s := q(x,x,u,u)$ (no sum on i).

We consider this last example in more detail, since it gives the class of all skew tridiagonal (i.e., bandwidth 1) finite differences.

Firstly, suppose we have the non-constant operator $\mathcal{J} = u\partial_x + \partial_x u$. Eq. (5.25) discretizes this if $s(x,u) = u$. The only constraint is the the tensor s_{ij} must be symmetric, to maintain the overall skew-symmetry of the finite difference. For example, we can take $s_{ij} = (u_i + u_j)/2$ (only elements with $|i-j| = 1$ are actually used). This gives the finite difference tensor

$$K = \frac{1}{2}\begin{pmatrix} \ddots & & & & & \\ -u_0 - u_1 & 0 & u_1 + u_2 & & & \\ & -u_1 - u_2 & 0 & u_2 + u_3 & & \\ & & -u_2 - u_3 & 0 & u_3 + u_4 & \\ & & & & \ddots & \end{pmatrix}, \tag{5.26}$$

showing how the skew-symmetry is maintained. [1]

Secondly, suppose we wish to difference on an irregular grid, where the data are known at the points $c(x_i) = c(ih)$—we know $v(c(x_i))$. Then we want an approximation of $v' = v_c$. From the chain rule, this is equal to v_x/c_x. We cannot get this by applying (5.25) to v, since the terms in v only cannot cancel.

We apply (5.25) to a function $w(x, v(c(x))$. This gives the equation

$$(s\partial_x + \partial_x s)w(x, v(c(x))) = (\sqrt{2s}\partial_x\sqrt{2s})w(x, v(c(x)) = \frac{v_x(c(x))}{c'(x)}$$

[1] Interestingly, the choice $s_{ij} = \sqrt{u_i u_j}$ actually gives a *Poisson* K. This is because it is the image of the standard central difference under a change of variables which sends ∂_x into $u\partial_x + \partial_x u$.

with solution

$$s = \frac{1}{2c'^2}, \quad w = c'v.$$

For the discretization, any w_j and any symmetric s_{ij} with this continuous limit can be taken. If c' is known analytically, we can use the midpoints of the intervals to get

$$s_{0,1} = \frac{1}{2(c'_{\frac{1}{2}})^2}, \quad s_{0,-1} = \frac{-1}{2(c'_{-\frac{1}{2}})^2},$$

$$v'_0 = \frac{1}{4h}\left(\frac{c'_1 v_1}{{c'_{\frac{1}{2}}}^2} - \frac{c'_{-1}v_{-1}}{{c'_{-\frac{1}{2}}}^2}\right) + \mathcal{O}(h^2)$$

—that is, a second order, 2-point, anti-symmetric discretization of the derivative on smooth grids. With constant spacing, $c(x) = x$ and it reduces to the standard central difference.

If c' is not known, we can approximate it symmetrically with

$$f_{0i} = \frac{1}{2}\left(\frac{c_i - c_0}{h}\right)^{-2},$$

$$w_i = \frac{(v_i + v_{i+1})(c_{i+1} - c_i) + (v_i + v_{i-1})(c_i - c_{i-1})}{4h},$$

which has bandwidth 2. It is not consistent on rough grids, however; the error is $\mathcal{O}(h^2 c'''') = \mathcal{O}(h^{-1})$ near a discontinuity of c'.

With $m > 1$ components and bandwidth 1, the two possibilities are $\boldsymbol{l} = (0,0)$ and $\boldsymbol{l} = (0,1)$. In the first case, the tensor T must be skew symmetric in its only two slots, but there are no group symmetries. In the second case, $G_{\boldsymbol{l}}$ has two elements, the identity and left translation. Under left translation, $(0,1) \mapsto (-1,0)$; applying $\sigma = (01)$ of sign -1 maps $(-1,0) \mapsto (0,-1)$. In the last step we apply the σ^{-1} to the indices of T. Combining both possibilities gives the finite difference

$$-T^{\mathrm{T}}v_{-1} + Jv_0 + Tv_1 = (J + T - T^{\mathrm{T}})v + h(T + T^{\mathrm{T}})v_x + \mathcal{O}(h^2),$$

where $J = -J^{\mathrm{T}}$. Further imposing the symmetry $i \to -i$, of sign -1, is equivalent to taking $T = T^{\mathrm{T}}$; then the finite difference is second order. As above, nonconstant operators are approximated by taking $T_l = T(x_l, x_{l+1}, u_l, u_{l+1})$, symmetric in its first and second pairs of arguments, and $J_l = J(x_{l-1}, x_l, x_{l+1}, u_{l-1}, u_l, u_{l+1})$, symmetric under (13) and (46).

Case 2. $p = 2$ **integrals,** $d = 1$ **space dimension.** With $m = 1$ component, the simplest tensor has base index $\boldsymbol{i} = (0,1,2)$. This will lead to bandwidth 2. We start with the arrow $1 \to 2$ (see Figure 1). Translating left by 1 and rotating indices right (an even permutation) (i.e., performing $(0,1,2) \mapsto (-1,0,1) \mapsto (0,1,-1)$) gives the arrow $1 \to -1$. Repeating gives the arrow $-2 \to -1$. The resulting diagram is already symmetric under $h \to -h$, so we do not need to add this operation. The

(a) (b)

FIGURE 1. Case 2, Two integrals, one space dimension. (a) One variable; (b) m variables. See Eq. (5.27) for the finite difference interpretation of (a).

(a) (b)

FIGURE 2. Case 2, Two integrals, one space dimension, m variables, $\boldsymbol{l} = (0, 0, 1)$.

diagram in Fig. 1(a) corresponds to the finite difference

$$\begin{aligned} F(0,1,2) &= \det V_{1,2} + \det V_{1,-1} + \det V_{-2,-1} \\ &= (v_1 w_2 - v_2 w_1) + (v_1 w_{-1} - v_{-1} w_1) + (v_{-2} w_{-1} - v_{-1} w_{-2}) \end{aligned} \tag{5.27}$$

Expanding in Taylor series, this is

$$h^3(3(v'w'' - w'v'') + 2(vw''' - wv''')) + \mathcal{O}(h^5)$$

With $m > 1$ component there is essentially one finite difference each with bandwidth 0, 1, and 2. We write $T_{\alpha_1,\alpha_2,\alpha_3} = T_{abc}$.

With $\boldsymbol{l} = (0, 0, 0)$ (bandwidth 0), Eq. (4.23) says we must have T_{abc} completely skew-symmetric. One might not call this a "difference," since it only acts on v_0 and w_0.

With $\boldsymbol{l} = (0, 0, 1)$ (bandwidth 1), Eq. (4.23) says we have $T_{abc} = -T_{bac}$. The weight of $\boldsymbol{l}$ is $n(\boldsymbol{l}) = 2$, since it has 2 zeros. $G_{\boldsymbol{l}}$ has two elements, the identity and a left translation. Applying the left translation followed by a shift-right permutation σ (of sign 1), $(0, 0, 1) \mapsto (-1, -1, 0) \mapsto (0, -1, -1)$, giving the arrow $-1 \to -1$ with label $\sigma^{-1}(abc) = bca$. $(0, -1, -1)$ has one zero, so the weight of this arrow is 1. Together we get two arrows, with diagram Fig. 2(a) and finite difference

$$\begin{aligned} &2T_{abc}(v_{0,b} w_{1,c} - v_{1,c} w_{0,b}) + T_{bca}(v_{-1,b} w_{-1,c} - v_{-1,c} w_{-1,b}) \\ &= (T_{abc} + T_{bca} + T_{cab}) v_b w_c + h(T_{abc} - T_{bca}) v_b w_c' + h(T_{cab} - T_{bca}) v_b' w_c + \mathcal{O}(h^2). \end{aligned}$$

(The initial, skew-symmetric term could be removed by a term $F(0, 0, 0)$.)

The resulting tensor is not invariant under $g : i \mapsto -i$. Its image under g is shown in Fig. 2(b). These two diagrams can be added or subtracted to get a tensor that is g-invariant with sign 1 or -1, as desired.

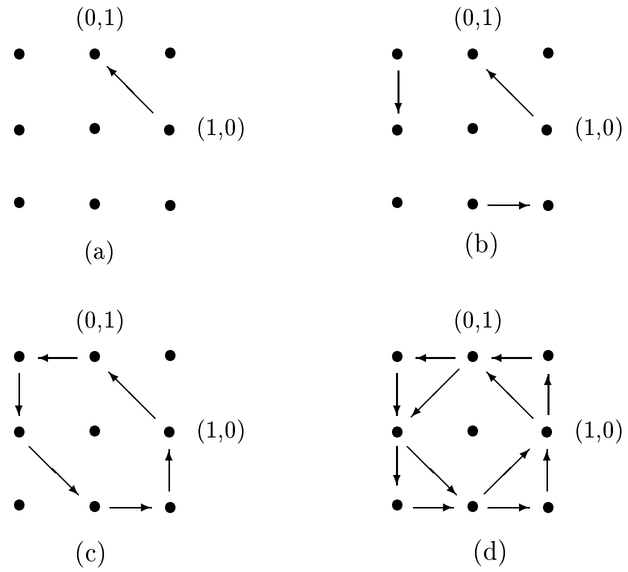

FIGURE 3. Case 4, Step-by-step construction of the Arakawa Jacobian.

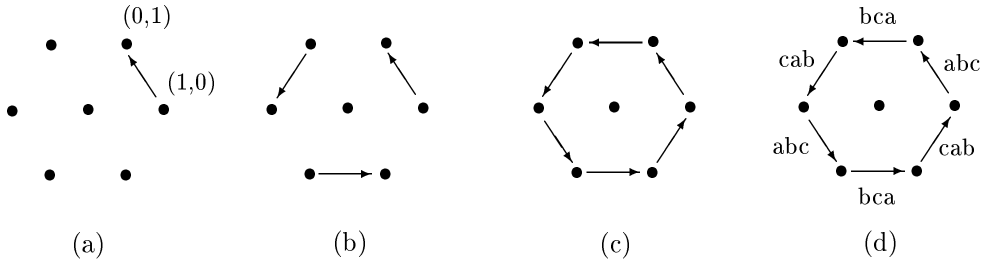

FIGURE 4. Case 4, The Arakawa Jacobian on a triangular grid.

With $\boldsymbol{l} = (0, 1, 2)$ (bandwidth 2), Eq. (4.23) says that T is arbitrary. The group $G_{\boldsymbol{l}}$ has three elements: the identity, and a shift left by 1 or 2. Apply the two translations gives the diagram Fig. 1(b). However, unlike this case with $m = 1$, this is not invariant under $i \mapsto -i$, i.e., it does not give a second-order finite difference. Applying this symmetry gives the second row of labels in Fig. 3(b). (For example, under $(0, 1, 2) \mapsto (0, -1, -2)$ the arrow $1 \to 2(abc)$ maps to the arrow $-1 \to -2(abc)$ with sign -1, or $-2 \to -1(abc)$ with sign 1. Adding these makes $T_{abc} = T_{bca}$, i.e., we can take T to be symmetric under even permutations.

Case 3. $p = 1$ **integral,** $d = 2$ **space dimensions.** With p free indices in K we can only couple unknowns which span a p-dimensional subspace of $\mathbb{R}^d$. This is equivalent to the case $d = p$. For example, on a square grid in $\mathbb{R}^2$, $F((0, 0), (0, 1)) = hv_y + \mathcal{O}(h^3)$.

Thus, to get fundamentally new finite difference tensors, we need $p \geq d$.

Case 4. $p = 2$ **integrals,** $d = 2$ **space dimensions.** Consider $m = 1$ and a square grid. The simplest index set $\boldsymbol{i}$ we can take is $\boldsymbol{i} = ((0, 0), (1, 0), (0, 1))$, as shown in

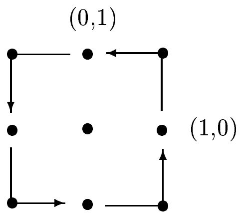

FIGURE 5. Case 5, 3 integrals in 2 dimensions

Fig. 3(a). Unfortunately, this has lattice bandwidth 2 and Euclidean bandwidth $\sqrt{2}$, an unavoidable property of the lattice.

Applying the two translations gives the diagram 3(b): the simplest translation-invariant skew tensor. It gives $h(v_x w_y - v_y w_x) + \mathcal{O}(h^2)$. That is, it is an "Arakawa"-type Jacobian.

With even operators on this grid, reflections (necessary for second order accuracy) and rotations coincide, which reduces the complexity of the generated finite difference. Applying them gives the diagram 3(c), a second-order Jacobian. Finally, applying the rotations by $\pi/2$ gives the diagram 3(d), a Jacobian with the full symmetry group $D_4 \circledS \mathbb{Z}^2$. As can be shown by expanding the entire finite difference, *Fig. 3(d) is the Arakawa Jacobian* (first derived in [1].)

We could have stopped at Fig. 3(c); its anisotropy may be irrelevant for some problems, and its complexity is half that of 2(d)—12 terms instead of 24.

Consider the same problem on a regular triangular grid. Now $\boldsymbol{i} = ((0,0),(1,0),(0,1))$ (Fig. 4(a)) will give a graph bandwidth of 1, not 2. Applying the two translations gives Fig. 4(b), and reflections Fig. 4(c), which *already* has the full symmetry of the grid. Thus Arakawa-type Jacobians are naturally suited to triangular grids. (Notice that Figs. 4(c) and 3(c) are essentially the same.)

There are two points to learn from this:

1. With p integrals, grids with $p+1$ mutual nearest neighbours around a cell will give tensors of bandwidth 1. This is only possible in dimension $d \geq p$.
2. On some grids, the (optional) spatial symmetries coincide with some of the (required) skew symmetries and/or reflection symmetry (needed for second order accuracy).

Cases 5 and 6 illustrate these points.

(An m-variable analogue of the Arakawa Jacobian is shown in Fig. 4(c). It approximates a complicated second order operator, but if T_{abc} is symmetric under even permutations, it is $3\sqrt{3}h^2 T_{abc}\mathcal{J}(v_b, w_c) + \mathcal{O}(h^4)$.)

Cases 5 & 6: $p = 3$ integrals in 2 & 3 dimensions. The above observation suggests that in two dimensions, the square grid, with 4 vertices around each cell, is better suited to the case of three rather than two integrals. With $\boldsymbol{i} = ((0,0),(1,0),(1,1),(0,1))$, applying the 3 translations only gives a tensor which is D_4-symmetric (Fig. 5). It equals

$$4h^2(v^1\mathcal{J}(v^2,v^3) + v^2\mathcal{J}(v^3,v^1) + v^3\mathcal{J}(v^1,v^2)) + \mathcal{O}(h^4)$$

where $\mathcal{J}$ is the Jacobian. Taking $v^3 = 1$, for example, recovers the Arakawa Jacobian Fig. 2(d), and shows that the Arakawa Jacobian also has the (Casimir) integral $\sum_i u_i$.

It also suggests that in three dimensions with three integrals, a face-centered-cubic grid (the red points in a red-black colouring of a cubic grid) is suitable. Each vertex is surrounded by 8 tetrahedra. Taking $\boldsymbol{l} = ((0,0,0),(0,1,1),(1,0,1),(1,1,0))$ (i.e, coupling the unknowns around one of the tetrahedra) leads to a fully symmetric discretization of the three-dimensional Jacobian $\det(\partial v^i/\partial x_j)$. Using a cubic grid with $\boldsymbol{l} = ((0,0,0),(0,0,1),(0,1,0),(1,0,0))$ leads to a 3D Jacobian with twice the complexity.

6. Discussion

We have presented a systematic method for discretizing PDEs with a known list of integrals. Since all vector fields f_i with integrals $I^1,\ldots,I^p$ can be written in the form (2.11), the vector fields $F(\boldsymbol{i})$ span all integral-preserving discretizations. The required symmetry properties of K make the the finite differences unavoidably complicated, but sometimes the (optional) spatial symmetries G coincide with the (compulsory) skew symmetries S_{p+1}, reducing the overall complexity of the finite difference.

We close with some comments on future directions.

1. We have not yet mentioned time integration. It may not as crucial to preserve integrals in time as in space; this is not usually done with the Arakawa Jacobian, for example. If it is important, we note that linear integrals are preserved by an consistent linearly covariant method (such as the Euler method used in (1.2)); quadratic integrals are preserved by some Runge-Kutta methods such as the midpoint rule; and any number of arbitrary integrals can be preserved by a discrete-time analogue of (3.12) [13]. With one integral, a simple method is based on splitting K [12].
2. To get simpler finite differences, some of the spatial symmetries can be broken, as for example in the half- and quarter-size Arakawa Jacobians in Fig. 3(b,c). How important is this in practice? These Jacobians are still fully translation invariant. Breaking *this* symmetry gives even simpler tensors K, in which, e.g., different differences are applied to red and to black points. Is this useful?
3. Such broken symmetries may be partially repaired "on the fly" during the time integration. At the nth time step we use the finite difference tensor $\mathrm{sgn}(g_n)gK$, with g_n ranging over the symmetries. This decreases the symmetry errors by one power of the time step [8], which, with $\Delta t = (\Delta x)^r$, may be plenty. Most drastically, K could be only first-order accurate, improving to second through the time integration. This would require a careful stability analysis.
4. Although our discretizations are not Hamiltonian (unless $\mathcal{K}$ is constant), they can be volume-preserving. The system (2.11) is volume preserving for all I^j if $\sum_i \partial K_{ijk\ldots}/\partial u_i = 0$ for all j, $k,\ldots$. In simple cases this is simple to arrange; K in Eq. (3.14) is not volume-preserving, but K in Eq. (5.26) is. Incorporating volume-preservation in general is more difficult; see [12] for a discussion.

5. We have deliberately avoid mentioning boundaries and the precise degree of smoothness required of the arguments that make $\mathcal{K}$ skew. These are studied in [14]. If the PDE develops shocks, a careful weighting of the $F(\boldsymbol{i})$ will be required to capture them well, the analogue of the many methods for choosing H in (1.2) [3]. The present work applies to infinite or (trivially) to periodic domains. With finite domains, one can start with $K_{\boldsymbol{i}}$ at an interior point $\boldsymbol{i}$, and extend it to the boundary by skew-symmetry, giving finite difference tensors satisfying certain "natural" boundary conditions. Also on finite domains, there is the possibility of using global (e.g. spectral) methods. Of course, these are in the span of our basis, but that is not the best way to view them. Preserving integrals with global methods is studied in [14].
6. We have concentrated on constructing skew-symmetric tensors approximating skew operators. Exactly the same technique can be used to construct symmetric tensors. We replace the canonical sign function on S_{p+1} by any sign function σ that makes $S_{p+1} \cong \mathbb{Z}_2$. If $\sigma(\boldsymbol{i}) = 1$ for all $\boldsymbol{i}$, for example, the resulting K is completely symmetric, and when contracted against any $p-1$ of the I^j, has real eigenvalues. If negative definite, the I^j decrease in time. What is the relationship with the support operator method [20]?

Acknowledgements. This paper had a long gestation, during which the support of the Isaac Newton Institute, Cambridge, and the MSRI, Berkeley, were invaluable. Useful discussions with Phil Morrison, Reinout Quispel, Nicolas Robidoux, and Rick Salmon are gratefully acknowledged. This work was supported in part by a grant from the Marsden Fund of the Royal Society of New Zealand.

References

[1] A. Arakawa, Computational design for long-term numerical integration of the equations of fluid motion: two-dimensional incompressible flow. Part I. *J. Comput. Phys.* **1** (1966) 119–143.
[2] R. W. R. Darling, *Differential forms and connections*, CUP, 1994.
[3] Arieh Iserles, *A First Course in the Numerical Analysis of Differential Equations*, Cambridge University Press, 1996.
[4] Z. Fei and L. Vasquez, Two energy-conserving numerical schemes for the Sine-Gordon equation, *App. Math. Comp.* **45**(1) (1991), 17–30.
[5] P. Gautheron, Some remarks concerning Nambu mechanics, *Lett. Math. Phys.* **37** (1996) 103–116.
[6] R.T. Glassey, Convergence of an energy-preserving scheme for the Zakharov equations in one space dimension, *Math. Comp.* **58**(197) (1992), 83–102
[7] R.T. Glassey, and J. Schaeffer, Convergence of a second-order scheme for semilinear hyperbolic equations in $2+1$ dimensions, *Math. Comp.* **56**(193) (1991), p. 87
[8] A. Iserles, R.I. McLachlan, and A. Zanna, Approximately preserving symmetries in the numerical integration of ordinary differential equations, *Eur. J. Appl. Math.*, to appear.
[9] P. Lax and B. Wendroff, Systems of conservation laws, *Comm. Pure Appl. Math.* **13** (1960), 217–237.
[10] D. Lewis, and J.C. Simo, Conserving algorithms for the dynamics of Hamiltonian systems on Lie groups, *J. Nonlinear Sci.* **4** (1994), 253–299.
[11] R.I. McLachlan, Symplectic integration of Hamiltonian wave equations, *Numer. Math.* **66** (1994), 465–492.

[12] R.I. McLachlan and G.R.W. Quispel, Generating functions for dynamical systems with symmetries, integrals, and differential invariants, *Physica D* **112** (1998) 298–309.
[13] R.I. McLachlan, G.R.W. Quispel, and N. Robidoux, Geometric integration using discrete gradients, *Phil. Trans. Roy. Soc. A*, to appear.
[14] R.I. McLachlan and N. Robidoux, Integral-preserving and skew-adjoint discretizations of partial differential equations, in preparation.
[15] Y. Nambu, Generalized Hamiltonian dynamics, Phys. Rev. D **7**, 2405–2412 (1973)
[16] P.J. Morrison, Princeton Plasma Physics Laboratory Report PPL–1783 (Princeton, New Jersey, 1981).
[17] Peter J. Olver, *Applications of Lie groups to differential equations*, Springer-Verlag, New York, 1986.
[18] G.R.W. Quispel and H.W. Capel, *Phys. Lett. A* **218** (1996), 223-228.
[19] R. Salmon and L. D. Talley, Generalizations of Arakawa's Jacobian, *J. Comput. Phys.* **83** (1989) 247–259.
[20] Mikhail Shashkov, *Conservative finite-difference methods on general grids*, CRC Press, Boca Raton, 1996.
[21] J.C. Simo, N. Tarnow, and K.K. Wong, Exact energy-momentum conserving algorithms and symplectic schemes for nonlinear dynamics, *Comp. Meth. Appl. Mech. Eng.* **100** (1992), 63–116.
[22] Michael Tabor and Yvain M. Treve, eds., *Mathematical Methods in Hydrodynamics and Integrability in Dynamical Systems*, AIP conference proceedings 88, AIP, 1982.
[23] L. Takhtajan, On Foundation of the Generalized Nambu Mechanics, *Commun. Math. Phys.* **160**, 295–315 (1994).
[24] V. Zeitlin, Finite-mode analogues of 2D ideal hydrodynamics: Coadjoint orbits and local canonical structure, *Physica D* **49** (1991), 353–362.

Appendix

Consider the operator $\mathcal{K} = \partial_x$. We were puzzled by the following: the tensor K in Eq. (3.13),

$$K = \frac{1}{2h}\begin{pmatrix} \ddots & & & & \\ -1 & 0 & 1 & & \\ & -1 & 0 & 1 & \\ & & & & \ddots \end{pmatrix},$$

preserves not just the integral H it operates on, but also $C = \sum_i u_i$, because C is a Casimir of K. But this two-integral discretization does not arise from any of the rank 3 skew-tensors we derived in Section 5, Case 2—Eq. (5.27) in particular. Contracting with the required integral C gives a discretization of ∂_{xxx}, not of ∂_x. The same is true for any other basis element.

To force C to appear explicitly in the discretization, we first find a skew differential operator $\mathcal{J}(u,v,w)$ such that $\mathcal{J}(u,v,\delta C) = \mathcal{K}(u,v)$ for all $\mathcal{H}$. If we restrict to a finite domain D so that $\int_D 1\,dx$ is finite, a natural solution is

$$\mathcal{J}(u,v,w) = uv_x\int_D w\,dx - uw_x\int_D v\,dx + u\int_D vw_x\,dx.$$

This is a non-local differential operator, which is the resolution of the paradox. It only reduces to a local operator when $w \equiv 1$. Discretizing its derivatives by central derivatives, and integrals $\int w\,dx$ by $\sum_i w_i$, gives a non-local skew 3-tensor J such that $J(\nabla C) = K$. In Section 5 we only looked at local tensors.

It seems unlikely that the telescoping sum which makes this example work will work for nonlinear Casimirs. On the other hand, it is quite hard to destroy linear ones. Therefore we suggest the following strategy: temporarily disregard any known linear integrals (mass, momentum etc.). Construct a skew tensor so as to preserve the desired nonlinear integrals. Then, check that this tensor has (some discretization of) the required linear integrals as Casimirs.

The situation is analogous to preserving volume, a linear *differential* invariant discussed in Section 5, note 4.

Mathematics, Massey University, Private Bag 11–222, Palmerston North, New Zealand
E-mail address: `R.McLachlan@massey.ac.nz`