Spatial discretization of partial differential equations with integrals

Robert I. McLachlan
IFS, Massey University, Palmerston North, New Zealand

E-mail: r.mclachlan@massey.ac.nz

We consider the problem of constructing spatial finite difference approximations on an arbitrary fixed grid which preserve any number of integrals of the partial differential equation and preserve some of its symmetries. A basis for the space of of such finite difference operators is constructed; most cases of interest involve a single such basis element. (The "Arakawa" Jacobian is such an element, as are discretizations satisfying "summation by parts" identities.) We show how the grid, its symmetries, and the differential operator interact to affect the complexity of the finite difference.

Key Words: finite differences, conservative, integrals, summation by parts

1. CONSERVATIVE DISCRETIZATION

Consider the following four approaches to constructing finite difference schemes that inherit some form of stability or conservation from the continuous problem.

1. For the conservation law

$$\dot{u} = \partial_x(f(u)), \tag{1}$$

we have the classic "conservative" discretizations

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} = \frac{H(u_{i+j}^n, \dots, u_{i-j+1}^n) - H(u_{i+j-1}^n, \dots, u_{i-j}^n)}{\Delta x} \tag{2}$$

introduced by Lax and Wendroff [13]. (Here $\dot{u} = \frac{\partial u}{\partial t}$, $\partial_x = \frac{\partial}{\partial x}$, and $u_i^n \approx u(i\Delta x, n\Delta t)$.) For any H, the semi-discretization (2) preserves a discrete total mass (a linear function), namely $\sum_i u_i^{n+1} = \sum_i u_i^n$.

2. The widely-used Arakawa Jacobian [1] is 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 positive quadratic functions. This gives the method a (semi-discrete) nonlinear stability. It has been generalized by Salmon and Talley [20].

3. In one space dimension, let $v = (v_0, \dots, v_n)^t$ be a grid function and let D be a differentiation matrix approximating $\partial/\partial x$. If D satisfies the "summation by parts" identity

$$\langle u, Dv\rangle = u_n v_n - u_0 v_0 - \langle Du, v\rangle$$

with respect to an inner product $\langle u, v\rangle = u^t S v$, then D can be used to construct stable finite difference methods, ones that preserve or decrease $\langle u, u\rangle$. Such non-Euclidean inner products were introduced by Kreiss and Scherer [12], and the approach has been recently developed further by Olsson [18] and by Carpenter et al. [5, 2] (see also [3] for a similar approach in the context of spectral methods).

4. Many schemes have been derived for particular PDEs (e.g. nonlinear wave equations) that preserve nonlinear, even non-quadratic, energy integrals [7, 8, 9].

In this paper we show that all four approaches are essentially equivalent, and give a complete local description of all such finite difference methods. The summation-by-parts property is not only sufficient, it is also necessary if the finite difference is to have a conservation law; moreover, not only a norm $\langle u, u\rangle$, but any (e.g. energy) integral, and any number of integrals, can be controlled in this way. However, extending the summation-by-parts property to systems in any number of variables and any number of space dimensions, and with more than one integral, introduces intriguing restrictions on the finite differences, which we explore in this paper.

We consider PDEs with independent spatial variables $x \in \mathbb{R}^d$, fields u belonging to a Banach space $\mathfrak{B}$ with values $u(x) \in \mathbb{R}^m$, and PDEs of the form

$$\dot{u} = \mathcal{D}(u)\frac{\delta \mathcal{H}}{\delta u} \tag{3}$$

where $\mathcal{D}(u)$ is a linear differential operator and $\mathcal{H}\colon \mathfrak{B} \to \mathbb{R}$ has a variational derivative in the sense that

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathcal{H}(u+\varepsilon v)|_{\varepsilon=0} = \int \frac{\delta H}{\delta u} v \,\mathrm{d}x$$

for all $u, v \in \mathfrak{B}$. Typically $\mathfrak{B} = C^k(\mathbb{R}^d, \mathbb{R}^m)$. In particular, the conservation law (1) takes this form, with x, $u \in \mathbb{R}$ and

$$\mathcal{H} = \int F(u(x))\,\mathrm{d}x, \quad \mathcal{D} = \partial_x, \quad F' = f. \tag{4}$$

Hamiltonian systems are in the form (3), where $\mathcal{D}$ is the Hamiltonian operator and $\mathcal{H}$ the energy (Hamiltonian). The PDE (3) preserves energy if $\mathcal{D}$ is skew-adjoint with respect to the L_2 inner product, i.e.

$$\int u\mathcal{D}v\,\mathrm{d}x = -\int v\mathcal{D}u\,\mathrm{d}x,$$

which we assume from now on.

The system (3) has $\mathcal{I}\colon \mathfrak{B} \to \mathbb{R}$ as an integral if $\dot{\mathcal{I}} = \int \frac{\delta I}{\delta u}\mathcal{D}(u)\frac{\delta \mathcal{H}}{\delta u}\,\mathrm{d}x = 0$. Some integrals $\mathcal{C}$ are distinguished in that $\mathcal{D}(u)\frac{\delta C}{\delta u} \equiv 0$; they are called Casimirs. The operator ∂_x has a single Casimir, $\mathcal{C} = \int u\,\mathrm{d}x$, because $\partial_x(\delta\mathcal{C}/\delta u) = \partial_x 1 = 0$.

We shall seek spatial discretizations of "skew-gradient" form

$$\dot{u} = A(u)\nabla H(u), \quad A^t = -A, \tag{5}$$

where now $u \in \mathbb{R}^n$, for this makes the "summation by parts" property manifest:

$$\dot{H} = (\nabla H)^t A \nabla H = 0.$$

However, much more is true: given *any* discretization H of the continuum energy $\mathcal{H}$, an arbitrary system $\dot{u} = f(u)$ has H as an integral if and only if it can be written in the form (5) [16, 19]. Therefore, the form (5) includes all possible conservative discretizations.

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

$$\dot{u}_{i_0} = f_{i_0} = \sum_{i_1,\ldots,i_p} K_{i_0\ldots i_p} \frac{\partial I^1}{\partial u_{i_1}} \cdots \frac{\partial I^p}{\partial u_{i_p}}. \tag{6}$$

We write (6) as

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

What about Casimirs? Suppose that instead of contracting K against all the integrals, as in (6), 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 (6)) will have I^1 as an integral. But there are many different Ks, 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 is present for Hamiltonian systems, disappears in this formulation. We are free to move between different representations of f as needed.

Looking back at Eq. (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 $\int u \, \mathrm{d}x$ is an integral, and the conservative scheme (2) maintains a discrete analog of this. Secondly, if f is a variational derivative, $f = \delta\mathcal{H}/\delta u$ say, then the skew-adjointness of ∂_x means that $\mathcal{H}$ is an integral. To preserve this property requires a discretization of the form (5). Note that the matrix A need not *a priori* have $\sum u_i$ as a Casimir, and, similarly, form (2) does not preserve any discrete energy. Thus, the two expressions of conservation laws are in fact independent.

In this paper we generalize the second form. That is, faced with Eq. (1) we will preserve the skew-adjointness of ∂_x, not its Casimir. The first form is deceptively simple in this example, because the Casimir is so simple. It is not clear how to modify (2) to incorporate different Casimirs. In the second form the integral appears explicitly and, once an antisymmetric A approximating ∂_x is found, *any* quantity can be conserved.

Before continuing, we mention one trivial but complete solution to the whole problem of constructing finite differences which preserve conserved quantities. 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 $\{f \cdot \nabla I^j = 0 \colon j = 1, \dots, p\}$. 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 couple all of the f_i. By incorporating more of the known structure of the problem we can work more systematically.

Outline of the paper.

In this paper we study the entire class of systems (6) in their own right, constructing elementary tensors $K_{i_0 \dots i_p}$. That is, we establish (in a sense defined below; see Eq. (8) and the paragraph before it)

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

where h is the grid spacing. 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 by a numerical quadrature, say

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

giving the conservative system of ODEs

$$\dot{u} = K(\nabla I^1, \dots, \nabla I^p). \tag{7}$$

We shall see that a crucial ingredient is the symmetry group G of the grid. (For example, for a uniform grid in one space dimension, $G = \mathbb{Z} \times \mathbb{Z}_2$, the translations

and reflections; often one only requires invariance under some subgroup, say the translations only.) Not only is it desirable that a finite difference should inherit this symmetry of the grid, it also makes the construction of the finite difference tensors K relatively straightforward. For, as their skew symmetry already necessarily couples the values of K at different grid points together, so using G we can get further such information, enough to make an enumeration of the skew Ks possible. We shall also see that there is an opportunity for interactions between the skew and G-symmetries, which can make the finite differences more or less complicated.

Below we develop some requirements on the tensors K, and construct all the elementary ones, for various numbers of integrals, space dimensions, and grids. We consider regular grids based on, e.g., square or triangular lattices. Section 2 begins with the necessary definitions and outlines the strategy which is to be followed in constructing the finite differences. This is carried out in three stages, from the simplest (one dependent variable) to the most complicated (several independent variables with arbitrary coupling). The general result is stated in Proposition 2.2. We also introduce a diagram notation (see Figs. 1–5) which makes it easy to draw the diagrams for all skew finite differences and to write down the finite difference associated with each diagram.

In Section 3 this theory is applied to a sequence of problems of increasing complexity, by gradually increasing the number of spatial dimensions and the number of conserved quantities. In this way we develop a "library" of all skew-adjoint finite differences. The Arakawa Jacobian, for example, arises naturally in this sequence.

2. DEFINITIONS & THEORY

The fundamental objects are the *grid* L, the *index set* M, the *symmetry group* G, and the *finite difference tensor* K, which we now define.

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

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

The dimension of the phase space of the semi-discretization (7) is thus $|M|$, the number of elements in M. A *grid function* is a real function on M; for example, the system state is given by the grid function $u\colon M \to \mathbb{R}$. (Equivalently, $u \in \mathbb{R}^{|M|}$.) 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. A discretization of a PDE is a vector field on $\mathbb{R}^{|M|}$.

For simplicity, we only consider the interpretation of $u_{l,\alpha}$ in which $u_{l,\alpha} \approx u_\alpha(x_l)$. (Staggered grids and u_i representing other functions 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 (7) are denoted $v^1, \dots, v^p$. To reduce the number of indices, when $p = 1$ we denote the single such function by v, and when $p = 2$ we denote the two such functions by v and w.

A discretization of an operator $\mathcal{K}$ is a skew $(p+1)$-tensor $K \in \bigwedge^{p+1}(\mathbb{R}^{|M|})$. We need a notation which labels sets of $p+1$ indices in $M = L \times \{1, \dots, m\}$. We will

separate out the two factors in M (the grid points and the dependent variables) and write $\mathbf{l} = (l_0, \dots, l_p)$, $l_j \in L$, $\boldsymbol{\alpha} = (\alpha_0, \dots, \alpha_p)$, $\alpha_j \in \{1, \dots, m\}$. The skew tensor K is then indexed by $\mathbf{i} = (\mathbf{l}, \boldsymbol{\alpha})$, i.e., $K_{\mathbf{i}} = K_{i_0 \dots i_p} = K_{\mathbf{l},\boldsymbol{\alpha}} \in \mathbb{R}$. Since we are trying to *construct* such tensors, intermediate steps will also involve non-skew tensors, i.e., real functions on M^{p+1}.

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

$$\sum_{i_1,\dots,i_p \in L} K_{i_0 \dots i_p} v^1(x_{i_1}) \dots v^p(x_{i_p}) = \mathcal{K}(v^1, \dots, v^p)(x_{i_0}) + \mathcal{O}(h^r) \tag{8}$$

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

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 $K_{\mathbf{i}} = 0$ for all $\mathbf{i}$ 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." K is *local* if its bandwidth is independent of the number of grid points.

The grid L may be structured, like a square or triangular lattice, or unstructured. Let G be a symmetry group acting on $M = L \times \{1, \dots, m\}$. It acts on $u \colon M \to \mathbb{R}$ by $(G \cdot u)(i) = u(G \cdot i)$. We usually consider only *spatial symmetries*, those which are the identity in their second slot, i.e. $g(l, \alpha) = (h(l), \alpha)$. 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 discretization 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 $D_3 \circledS \mathbb{Z}^2$ as a symmetry group.

Some non-spatial 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.

We now investigate which Ks lead to G-invariant ODEs of the form (7).

Proposition 2.1. *Let G be a symmetry group acting on M. The ODE (7) is G-invariant if (i) the functions $I^1, \dots, I^p$ are G-invariant, (ii) K does not depend on u, and (iii)*

$$K_{g(i_0) \dots g(i_p)} = K_{i_0 \dots i_p}$$

for all $g \in G$ and for all $i_0, \dots, i_p$.

Proof. First let $h\colon \mathbb{R}^{|M|} \to \mathbb{R}^{|M|}$ be an arbitrary map of phase space. Under $w = h(u)$, the ODE (7) transforms to

$$\dot{w}_{i_0} = \sum_{\substack{i'_0,\ldots,i'_p \\ i_1,\ldots,i_p}} H_{i_0 i'_0} \widetilde{K}_{i'_0 \ldots i'_p} H_{i_1 i'_1} \frac{\partial \widetilde{I}^1}{\partial w_{i_1}} \cdots H_{i_p i'_p} \frac{\partial \widetilde{I}^p}{\partial w_{i_p}} \tag{9}$$

where $\widetilde{K}(w) = K(h^{-1}(w)) = K(u)$, $\widetilde{I}^j(w) = I^j(h^{-1}(w)) = I^j(u)$, and $H_{ij}(w) = \frac{\partial h_i}{\partial u_j}(h^{-1}(w))$. (In Eq. (9), the functions K, I^j, and H_{ij} are evaluated at w.) The ODE is h-invariant if

$$\dot{w}_{i_0} = \sum_{i_1,\ldots,i_p} K_{i_1 \ldots i_p} \frac{\partial I^1}{\partial w_{i_1}} \cdots \frac{\partial I^p}{\partial w_{i_p}}.$$

By assumption (i), $\widetilde{I}^j(w) = I^j(w)$ for all j. By assumption (ii), $\widetilde{K}(w) = K(w)$. For a symmetry h coming from a symmetry g which acts on M, we have $h_i(u) = u_{g(i)}$ and so $H_{ij} = \frac{\partial h_i}{\partial u_j}(h^{-1}(w)) = \delta_{j,g(i)}$. Eq. (9) then becomes

$$\begin{aligned}
\dot{w}_{i_0} &= \sum_{\substack{i'_0,\ldots,i'_p \\ i_1,\ldots,i_p}} \delta_{i'_0,g(i_0)} \cdots \delta_{i'_p,g(i_p)} K_{i'_0 \ldots i'_p} \frac{\partial I^1}{\partial w_{i_1}} \cdots \frac{\partial I^p}{\partial w_{i_p}} \\
&= \sum_{i_1,\ldots,i_p} K_{g(i_0)\ldots g(i_p)} \frac{\partial I^1}{\partial w_{i_1}} \cdots \frac{\partial I^1}{\partial w_{i_p}} \\
&= \sum_{i_1,\ldots,i_p} K_{i_0,\ldots,i_p} \frac{\partial I^1}{\partial w_{i_1}} \cdots \frac{\partial I^p}{\partial w_{i_p}},
\end{aligned}$$

where the last equality follows from assumption (iii), the symmetry of K. ∎

That is, if we extend the action of G on M to an action on $\bigwedge^{p+1}(\mathbb{R}^{|M|})$ by

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

which we write as

$$(gK)_{\mathbf{i}} = g(K_{\mathbf{i}}) = K_{g(\mathbf{i})},$$

and say the tensor K is *G-invariant* if $gK = K$ for all $g \in G$, then skew, G-invariant tensors K and invariant discretizations of the integrals I^j will lead to G-invariant discretizations of the form of Eq. (7).

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, $\operatorname{sgn}(g) = \pm 1$, 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. In this situation the tensor K is G-invariant if $gK = (\operatorname{sgn} g)K$ for all $g \in G$.

Thus we impose all the following requirements on the finite difference operator K:

1. K should be completely skew-symmetric;
2. K should be G-invariant;
3. K should be as simple as possible;
4. K should approximate the desired continuous operator to the desired order; and
5. K should be local.

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

$$\sum_{\sigma \in S_{p+1}, g \in G} \operatorname{sgn}(\sigma) \operatorname{sgn}(g) K_{\sigma(g(\mathbf{i}))} \tag{10}$$

is completely skew-symmetric and G-invariant. Here S_{p+1} is the symmetric group on $p+1$ elements. This suggests two ways to construct symmetric Ks:

1. Start with a K which approximates the desired continuous operator, and symmetrize it; or
2. 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 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 finite difference operators, from which linear combinations can be taken as desired. However, the form (10) 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 now derive finite differences in this form in three stages: firstly, for $m = 1$ dependent variable; secondly, for $m \geq 1$ dependent variables with no unknowns at the same grid point coupled; thirdly, the general case, $m \geq 1$ dependent variables with arbitrary coupling, culminating in the general statement, Proposition 2.2.

Case 1. $m = 1$ dependent variable

When $m = 1$ we drop the second component α of the index $i = (l, \alpha)$. By translation symmetry, we only need to calculate the finite difference tensor at a single grid point, which we will label 0. Fix a multi-index $\mathbf{i} \in L^{p+1}$ where $i_0 = 0$ and start with the elementary tensor defined by

$$K_{\mathbf{j}} = \begin{cases} 1, & \mathbf{j} = \mathbf{i}, \\ 0, & \mathbf{j} \neq \mathbf{i}. \end{cases}$$

That is, we start with a discrete delta function for K; the resulting finite difference at the point $i_0 = 0$ will couple, at a minimum, $v_{i_1}^1$, $v_{i_2}^2, \dots$, and $v_{i_p}^p$.

We assume that the indices in $\mathbf{i}$ 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

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

The finite difference associated to $H_{\mathbf{l}}$ at the grid point 0 is $\sum_{\mathbf{l}} H_{\mathbf{l}} v_{\mathbf{l}}$ where $l_0 = 0$. Since K is a discrete delta function, there is only one nonzero term in this sum, i.e.,

$$\sum_{\mathbf{l}} H_{\mathbf{l}} v_{\mathbf{l}} = \sum_{\rho, h, \mathbf{l}} \operatorname{sgn}(\rho) \operatorname{sgn}(h) v_{\mathbf{l}}, \tag{11}$$

where the sum is taken over all ρ, h, and $\mathbf{l}$ such that $\rho(h(\mathbf{l})) = \mathbf{i}$ and $l_0 = 0$. Therefore, $\mathbf{l} = \rho^{-1}(h^{-1}(\mathbf{i}))$. Let $g = h^{-1}$, $\sigma = \rho^{-1}$ so that Eq. (11) becomes

$$\sum_{\mathbf{l}} H_{\mathbf{l}} v_{\mathbf{l}} = \sum_{\sigma, g} \operatorname{sgn}(\sigma) \operatorname{sgn}(g) v_{\sigma(g(\mathbf{i}))} \tag{12}$$

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

This sum can be further simplified by finding the σs for each g. Recall that the indices in $\mathbf{i}$ are distinct. For each g in the sum (12), choose one corresponding σ such that $\sigma(g(\mathbf{i}))_0 = 0$. Then because the indices in $\mathbf{i}$ are distinct, the other permutations that satisfy this equation range over S_p, the permutations of the last p indices of $\sigma(g(\mathbf{i}))$. The sum over S_p can be evaluated to give a determinant, giving the finite difference at the grid point 0,

$$F(\mathbf{i}) := \sum_{g \in G_{\mathbf{i}}} \operatorname{sgn}(\sigma) \operatorname{sgn}(g) \det V_{\sigma(g(\mathbf{i}))} \tag{13}$$

where

$$G_{\mathbf{i}} = \{g \in G \colon \exists \sigma \in S_{p+1} \text{ such that } \sigma(g(\mathbf{i}))_0 = 0\}, \tag{14}$$

and the $p \times p$ matrix $V_{\mathbf{l}}$ has (j,k) entry

$$(V_{\mathbf{l}})_{jk} = v^j_{l_k}, \quad 1 \le j, k \le p. \tag{15}$$

(The first component of $\mathbf{l}$ is $l_0 = 0$ and does not appear in the matrix.) We call $G_{\mathbf{i}}$ the *symmetry set* of the index set $\mathbf{i}$.

We introduce a graphical notation for formulas such as Eq. (13). An arrow connecting grid points $i_1, \ldots, i_p$ will indicate a term $\det V_{\mathbf{i}}$. 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 (14) 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. (24) to see how easy this is.

Thus, in terms of the diagrams, constructing a skew finite difference amounts to choosing an initial arrow, finding the symmetry set $G_{\mathbf{i}}$, and finding the image of the initial arrow under $G_{\mathbf{i}}$.

Case 2. $m \geq 1$ dependent variables, distinct points coupled

Here "distinct points coupled" means that in the finite difference, no two dependent variables are coupled together at the *same* grid point, only at different grid points. (This case is introduced mostly as a stepping stone to Case 3.)

With this notation, "distinct points coupled" means that if $K_{\mathbf{l},\boldsymbol{\alpha}} \neq 0$, then $l_j \neq l_k$ for all $j \neq k$.

Suppose we were to start, as in Case 1, with a single index $\mathbf{i} = (\mathbf{l}, \boldsymbol{\alpha})$, with $l_0 = 0$, and to construct the symmetrized finite difference corresponding to the discrete delta function based at $\mathbf{i}$. We would find that the resulting finite difference would be, in general, non-zero at $(0, \boldsymbol{\beta})$ for all $\boldsymbol{\beta}$. To study the finite differences at grid point 0 it is best to instead construct, not a single basis finite difference, but the family of finite differences spanned by $K_{\mathbf{l},\boldsymbol{\alpha}}$ *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 $T \in \bigotimes^{p+1} \mathbb{R}^m$ and start with the (non-skew) tensor

$$K_{\mathbf{j},\boldsymbol{\alpha}} = \begin{cases} T_{\boldsymbol{\alpha}}, & \mathbf{j} = \mathbf{i}; \\ 0, & \mathbf{j} \neq \mathbf{i}. \end{cases}$$

Passing from Eq. (12) to Eq. (13) only required that the l_j be distinct; therefore the symmetrized finite difference at the point 0 (corresponding to Eq. (13) in the case of a single dependent variable) is the vector field

$$\sum_{\substack{g \in G_{\mathbf{i}} \\ \alpha_0, \ldots, \alpha_p}} \operatorname{sgn}(\sigma) \operatorname{sgn}(g) T_{\boldsymbol{\alpha}} \det V_{\sigma(g(\mathbf{i}))} \frac{\partial}{\partial u_{\sigma(g(\boldsymbol{\alpha}))_0}} \tag{16}$$

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 finite difference at grid point 0 in component α_0,

$$F(\mathbf{i}) := \sum_{\substack{g \in G_{\mathbf{i}} \\ \alpha_1, \ldots, \alpha_p}} \operatorname{sgn}(\sigma) \operatorname{sgn}(g) T_{g^{-1}(\sigma^{-1}(\boldsymbol{\alpha}))} \det V_{\sigma(g(\mathbf{l})),\boldsymbol{\alpha}} \tag{17}$$

Notice that in (17), 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 (17). To the arrow $\sigma(g(\mathbf{l}))$ we attach the label $g^{-1}(\sigma^{-1}(\boldsymbol{\alpha}))$ indicating the T-tensor attached to that determinant.

Case 3. $m \geq 1$ dependent variables, arbitrary coupling

This is the most general case, of which Cases 1 and 2 above are specializations.

Equality of some of the l_j affects the sum over permutations in Eq. (12). Let $\mathbf{j} = h(\rho(\mathbf{l}))$ where $j_0 = 0$. Let $n(\mathbf{j})$ be the number of 0s in $\mathbf{j}$. We need to determine the subgroup $S(\mathbf{j})$ of S_{p+1} leaving $\sigma(\mathbf{j})_0 = 0$. (It is not S_p as it was before.) First apply a permutation σ such that $\sigma(\mathbf{j})_k = 0$ for $k = 0, \ldots, n(\mathbf{j})$. Then any element of $S(\mathbf{j})$ can be constructed by first swapping the leading 0 with one of the other 0s, and then applying an arbitrary permutation of the last p elements. That is,

$$S(\mathbf{j}) = \{\sigma.(0k) \colon \sigma \in S_p,\ k = 1, \ldots, n(\mathbf{j}) - 1\}.$$

$S(\mathbf{j})$ has size $(p!)n(\mathbf{j})$.

Now note that $S(\mathbf{j})$ contains all permutations of the first $n(\mathbf{j})$ elements; summing over these merely skew-symmetrizes T with in these first $n(\mathbf{j})$ slots. We could have imposed this skew-symmetry in the first place without loss of generality. Since there will usually be g_j (a translation, say) with $g_j(l_j) = 0$, this skew-symmetry of T is true for any set of equal elements of $\mathbf{l}$.

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

$$\text{take } T \text{ to skew-symmetric under } \Sigma_{\mathbf{l}}. \tag{18}$$

With this assumption, each flip $(0k)$ in $S(\mathbf{j})$ gives an equal contribution, and we can evaluate the sum over permutations to give the symmetrized finite difference at the point 0, in component u_{α_0},

$$F(\mathbf{i}) := \sum_{\substack{g\in G_{\mathbf{i}} \\ \alpha_1,\dots,\alpha_p}} \operatorname{sgn}(\sigma)\operatorname{sgn}(g) n(g(\mathbf{l})) T_{g^{-1}(\sigma^{-1}(\boldsymbol{\alpha}))} \det V_{\sigma(g(\mathbf{l})),\boldsymbol{\alpha}}. \tag{19}$$

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

We summarize the above calculation in the following proposition.

PROPOSITION 2.2. *The finite differences $F(\mathbf{i})$, $\mathbf{i} = (\mathbf{l}, \boldsymbol{\alpha})$, $\mathbf{l} = (l_0, \dots, l_p)$, $\boldsymbol{\alpha} = (\alpha_0, \dots, \alpha_p)$, $(l_j, \alpha_j) \in M$, where $l_0 = 0$, given by Eqs. (19) together with (18), (14), and (15), span all G-invariant skew-symmetric finite differences at the point $(0, \alpha_0)$. Eq. (19) specializes to Eq. (17) when $l_j \neq l_k$ for all $j \neq k$, and specializes further to Eq. (13) when $m = 1$.*

In practice, 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.

3. THE LIBRARY OF SKEW-ADJOINT FINITE DIFFERENCES

Case 1. $p = 1$ integral, $d = 1$ space dimension.

We initially consider a constant, infinite grid, so as to get a large symmetry group $G = \mathbb{Z}$, the translations. By scaling it is sufficient to consider $\mathbf{i} = (0, 1)$. The bandwidth is 1. The symmetry set $G_{\mathbf{i}}$ contains only the identity and the left translation $i \mapsto i - 1$. The permutation which brings 0 to the front is $(0, -1) \mapsto (-1, 0)$, with sign -1. Thus we get the standard central difference

$$F(0, 1) = v_1 - v_{-1}. \tag{20}$$

On a grid with constant spacing h,

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

All other examples in this case must be linear combinations and/or scalings of 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. Given the skew tensor K, we can apply any linear operator to it which preserves the skew symmetry. (This must break the translation symmetry, for we have already found the most general translation-invariant K.) If we do not want the bandwidth to increase then the most general such combination is

$$K'_{ij} = S_{ij} K_{ij} \quad \text{(no sum on } i,j) \tag{21}$$

where S is a symmetric matrix. 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 finite difference $F(0,1)$, then

$$\sum_j K'_{ij} v_j = \left(h(s\partial_x + \partial_x s)v\right)|_{x=x_i} + \mathcal{O}(h^3), \tag{22}$$

where $s := q(x,x,u,u)$.

4. For example, suppose we have the non-constant operator $\mathcal{D}(u) = u\partial_x + \partial_x u$. Eq. (22) discretizes this if $s(x,u) = u$. The only constraint is that S must be symmetric, to maintain the overall skew-symmetry of the finite difference. For example, we can take $S_{ij} = (u_i + u_j)/(2h)$. This gives the finite difference tensor

$$K' = \frac{1}{2h}\begin{pmatrix} \ddots & & & & & \\ -u_0 - u_1 & 0 & u_1 + u_2 & & & \\ & -u_1 - u_2 & 0 & u_2 + u_3 & & \\ & & -u_2 - u_3 & 0 & u_3 + u_4 & \\ & & & & \ddots \end{pmatrix}, \tag{23}$$

showing how the skew-symmetry is maintained.[1] In the case $\mathcal{H}(u) = \int \frac{1}{2}u^2 \,\mathrm{d}x$, so that the PDE (3) is the inviscid Burgers' equation $\dot{u} = 3uu_x$, the semi-discretization $\dot{u} = K'\nabla H$, $H = \sum \frac{1}{2}u_i^2$ with K' given by Eq. (23) is

$$\dot{u}_i = \left((u_i + u_{i+1})u_{i+1} - (u_i + u_{i-1})u_{i-1}\right)/(2h).$$

This coincides with the well-known (e.g. [21]) result that the semi-discretization

$$\dot{u}_i = \frac{3\theta}{4h}(u_{i+1}^2 - u_{i-1}^2) + \frac{3(1-\theta)}{2h} u_i (u_{i+1} - u_{i-1})$$

conserves $\sum_i u_i^2$ only for $\theta = 2/3$.

[1] Interestingly, the choice $S_{ij} = \sqrt{u_i u_j}$ actually gives a *Poisson* K (see [14]). This is because K is then the image of the standard central difference approximation of ∂_x under a change of variables which sends ∂_x into $u\partial_x + \partial_x u$.

5. This approach also constructs differences on non-constant grids, using the mapping method. We then have to consider the approximation of the continuum conserved quantity $\mathcal{H}$ and its derivative. (We refer to [11, 17] for details.) Briefly, let the conserved quantity $\mathcal{H} = \int h(u)\,du$ be approximated by a quadrature $H = w^t h(u)$, where w is a vector of quadrature weights. Then $\delta\mathcal{H}/\delta u = h'(u) \approx W^{-1}\nabla H$, where $W = \mathrm{diag}(w)$. Therefore, the discrete approximation D of $\mathcal{D}$ must be approximated by an antisymmetric matrix times W; without loss of generality we can write

$$D = (W^{-1}KW^{-1})W = W^{-1}K.$$

That is, the skew factor has the form (21) with $S_{ij} = (w_i w_j)^{-1}$, and *only* this form is allowed for arbitrary (nonlinear) conserved quantities $\mathcal{H}$. We are forced to use approximations of $\mathcal{D}$ which are skew-adjoint (satisfy a summation-by-parts identity) with respect to a diagonal inner product W.[2] (If $\mathcal{H}$ is quadratic, then non-diagonal inner products $H = u^t W u$ can be used.)

To apply the mapping method, suppose the data are known at the points $x(c_i) = x(ih)$; we want an approximation of v_x. From the chain rule, this is equal to v_c/x_c, so we let $W = \mathrm{diag}(x_c(c_i))$ and K be the matrix corresponding to the finite difference $F(0,1)$ (i.e., central differences). The antisymmetric matrix appearing in (5) is $A = W^{-1}KW^{-1}$.

If x_c is not known, we can approximate it, e.g. by $(x_{i+1} - x_{i-1})/(2h)$, leading to the skew-adjoint approximation

$$\frac{v_{i+1} - v_{i-1}}{x_{i+1} - x_{i-1}} = v_x + \frac{h^2}{6}(3x_{cc}v_{xx} + x_c^2 v_{xxx}) + \mathcal{O}(h^4).$$

Near a discontinuity of x_c, the order of this approximation drops from 2 to 1.

With $m > 1$ dependent variables, we must choose $\mathbf{l}$, the indices of the grid points which are to be coupled together. With bandwidth 1, the two possibilities are $\mathbf{l} = (0,0)$ and $\mathbf{l} = (0,1)$. In the first case, the tensor T must be skew symmetric in its only two slots, but the symmetry set $G_{\mathbf{l}}$ contains only the identity, because any nonzero translation leaves $0 \notin g(\mathbf{l})$. In the second case, $G_{\mathbf{l}}$ 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 σ^{-1} to the indices of T, which is the same as taking its transpose. Combining both possibilities gives the finite difference

$$-T^t v_{-1} + J v_0 + T v_1 = (J + T - T^t)v + h(T + T^t)v_x + \mathcal{O}(h^2),$$

where $J = -J^t$ and J and T are $m \times m$ matrices. Further imposing the reflection symmetry $l \mapsto -l$, of sign -1, is equivalent to taking $T = T^t$; then the finite difference is second order.

[2] If W is not diagonal, we can use the resulting Ds in nonlinear equations; but the discretization will not conserve energy. This is done for Burgers' equation in [5]. Whether this lack of conservation is important is another matter!

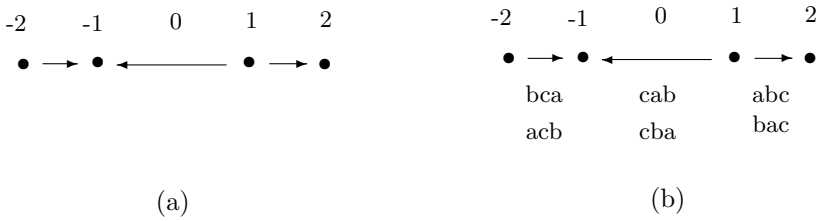

FIG. 1. Case 2, Two integrals, one space dimension. (a) One dependent variable; (b) m dependent variables. See Eq. (24) for the finite difference interpretation of (a).

FIG. 2. Case 2, Two integrals, one space dimension, m dependent variables. (a): $\mathbf{l} = (0,0,1)$. The circular arrow $-1 \to -1$ corresponds to the finite difference $\sum_{b,c} T_{bca} V_{(-1,-1),(b,c)} = \sum_{b,c} T_{bca}(v_{-1,b}w_{-1,c} - v_{-1,c}w_{-1,b})$. (b): The image of (a) under the reflection $l \mapsto -l$.

Case 2. $p = 2$ integrals, $d = 1$ space dimension

This example illustrates the construction of Case 2 of Section 2.

With $m = 1$ dependent variable, the simplest basis element has index $\mathbf{i} = (0,1,2)$. This will lead to bandwidth 2. The index $\mathbf{i} = (0,1,2)$ corresponds to 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 the reflection $l \mapsto -l$, an example of an interaction between the symmetries of the grid and of the finite difference tensor. The finite difference is automatically second-order accurate.

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

$$\begin{aligned} F(0,1,2) &= \det V_{1,2} + \det V_{1,-1} + \det V_{-2,-1} \\ &= (v_1 w_2 - v_2 w_1) + (v_1 w_{-1} - v_{-1} w_1) + (v_{-2} w_{-1} - v_{-1} w_{-2}) \end{aligned} \tag{24}$$

Expanding in Taylor series, this is

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

With $m > 1$ dependent variables there is essentially one finite difference each with bandwidth 0, 1, and 2. Their construction illustrates Case 3 of the Section 2. We write $T_{\alpha_0\alpha_1\alpha_2} = T_{abc}$.

With $\mathbf{l} = (0,0,0)$ (bandwidth 0), Eq. (18) 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 $\mathbf{l} = (0,0,1)$ (bandwidth 1), Eq. (18) says we must have $T_{abc} = -T_{bac}$. The weight of $\mathbf{l}$ is $n(\mathbf{l}) = 2$, since it contains 2 zeros. Thus, to the index $\mathbf{l}$ we associate

the arrow $0 \to 1$ with label abc(2). $G_{\mathbf{l}}$ has two elements, the identity and a left translation. Applying the left translation followed by a shift-right permutation σ (of sign 1), $(0,0,1) \mapsto (-1,-1,0) \mapsto (0,-1,-1)$, gives the arrow $-1 \to -1$ with label $\sigma^{-1}(abc) = bca$. $(0,-1,-1)$ contains one zero, so the weight of this arrow is 1. Together we get two arrows, with diagram Fig. 2(a) and finite difference (in the u_a-component) given by

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

(The initial, skew-symmetric term on the right hand side could be removed by a term $F(0,0,0)$.)

The resulting tensor is not invariant under $g\colon l \mapsto -l$. 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.

With $\mathbf{l} = (0,1,2)$ (bandwidth 2), Eq. (18) says that T is arbitrary. The symmetry set $G_{\mathbf{l}}$ 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 the reflection $l \mapsto -l$, 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 $-2 \to -1(acb)$ with sign 1.) As in the previous paragraph, these two diagrams can be added or subtracted to get a second-order finite difference.

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 $\mathbf{i}$ we can take is $\mathbf{i} = ((0,0),(1,0),(0,1))$, as shown in Fig. 3(a). Unfortunately, this has lattice bandwidth 2 and Euclidean bandwidth $\sqrt{2}$, an unavoidable property of the lattice.

We shall see that finite differences invariant under the full symmetry group of the lattice, namely $D_4 \circledS \mathbb{Z}^2$, are rather complicated. We begin instead with just $G = \mathbb{Z}^2$, the translations. The symmetry set $G_{\mathbf{i}}$ then contains three elements, namely translations by $(0,0)$, $(-1,0)$, and $(0,-1)$. Applying these 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.

Including rotations by π in G adds three more arrows to the diagram of the finite difference, shown in Fig. 3(c). Notice that it is also invariant under $l \to -l$, so

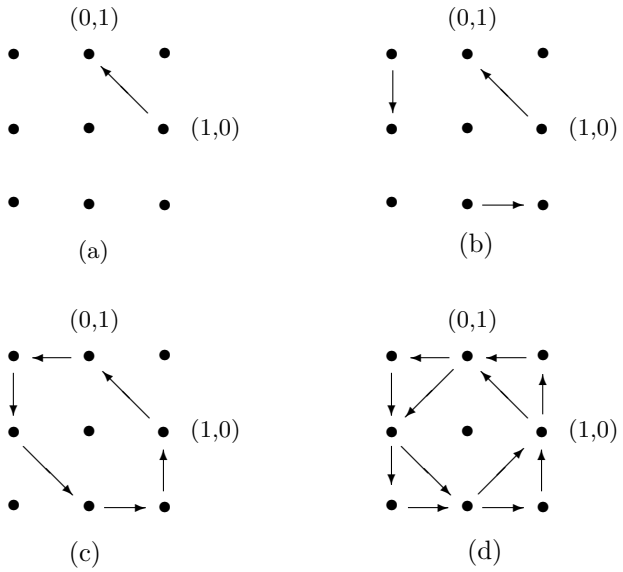

FIG. 3. Case 4, Step-by-step construction of the Arakawa Jacobian.

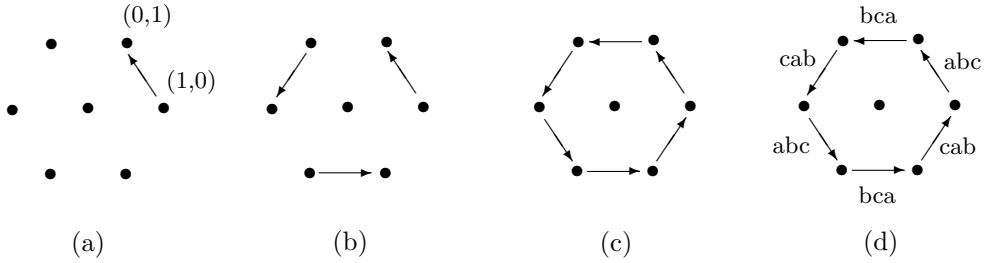

FIG. 4. Case 4, The Arakawa Jacobian on a triangular grid.

this is a second-order Jacobian—another example of a nice interaction between the grid and the required symmetries. Finally, including 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 $\mathbf{i} = ((0,0),(1,0),(0,1))$ (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:

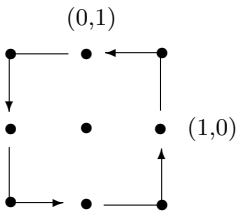

FIG. 5. Case 5, 3 integrals in 2 dimensions

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 further illustrate these points.

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

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 $\mathbf{i} = ((0,0),(1,0),(1,1),(0,1))$, 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 $\mathbf{l} = ((0,0,0),(0,1,1),(1,0,1),(1,1,0))$ (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 $\mathbf{l} = ((0,0,0),(0,0,1),(0,1,0),(1,0,0))$ leads to a three-dimensional Jacobian with twice the complexity.

4. A PUZZLE

Consider the operator $\mathcal{K} = \partial_x$. We were puzzled by the following: the tensor K corresponding to Eq. (20),

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

preserves not just the integral H whose gradient 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 3, Case 2—Eq. (24) 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 $\int_D 1 \, \mathrm{d}x$ is finite, a natural solution is

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

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 $\int w \, \mathrm{d}x$ by $\sum_i w_i$, gives a non-local skew 3-tensor J such that $J(\nabla C) = K$. In Section 3 we only looked at local tensors.

It seems unlikely that this example can be generalized to work for nonlinear Casimirs. On the other hand, it is quite hard to destroy linear Casmirs. 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.

5. 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 (6), the finite differences $F(\mathbf{i})$ 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 be 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

(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 (7) [16]. With one integral, a simple method is based on splitting K [15].

2. To get simpler finite differences, one can construct finite differences which are only invariant under a subgroup of the full symmetry group of the lattice or differential operator. See, for example, 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 [10], 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 (6) 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, e.g. K in Eq. (23) is volume preserving. Incorporating volume preservation in general is more difficult; see [15] for a discussion.

5. 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} homomorphic to $\mathbb{Z}_2$. If $\sigma(\mathbf{i}) = 1$ for all $\mathbf{i}$, for example, the resulting K is completely symmetric, and when contracted against any $p-1$ of the I^j, has real eigenvalues. If K is negative definite, then the I^j decrease in time for all I^j, which can be a useful property to preserve in difficult diffusion problems.

6. If the grid is rough, with no structure at all, then the constructions of Section 3 do not apply. Instead, we must directly construct approximations of $\mathcal{K}$ of the form $w_{i_0}^{-1} K_{i_0 \ldots i_p}$, where K is skew and w_i are quadrature weights (see Section 3, Case 1). We have studied such skew-adjoint approximations (in one space dimension, with one conserved quantity) on rough grids in [11].

ACKNOWLEDGMENT

The support of the Mathematical Sciences Research Institute, Berkeley, has been 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. MSRI wishes to acknowledge the support of the NSF through grant no. DMS–9701755.

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. M. H. Carpenter and J. Otto, High-order "cyclo-difference" techniques: An alternative to finite differences, *J. Comput. Phys.* **118** (1995), 242–260.
3. M. H. Carpenter and D. Gottlieb, Spectral methods on arbitrary grids, *J. Comput. Phys.* **129** (1996), 74–86.
4. R. W. R. Darling, *Differential forms and connections*, CUP, 1994.
5. M. H. Carpenter, J. Nordström, and D. Gottlieb, A stable and conservative interface treatment of arbitrary spatial accuracy, *J. Comput. Phys.* **148** (1999), 341–365.
6. A. Iserles, *A First Course in the Numerical Analysis of Differential Equations*, Cambridge University Press, 1996.
7. Z. Fei and L. Vasquez, Two energy-conserving numerical schemes for the Sine-Gordon equation, *App. Math. Comp.* **45**(1) (1991), 17–30.
8. R. T. Glassey, Convergence of an energy-preserving scheme for the Zakharov equations in one space dimension, *Math. Comp.* **58**(197) (1992), 83–102
9. 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
10. A. Iserles, R. I. McLachlan, and A. Zanna, Approximately preserving symmetries in the numerical integration of ordinary differential equations, *Eur. J. Appl. Math.* **10**(5) (1999), 419–445.
11. A. Kitson, R. I. McLachlan and N. Robidoux, Skew-adjoint finite difference methods on non-uniform grids, *New Zealand J. Math.*, to appear.
12. H.–O. Kreiss and G. Scherer, Finite element and finite difference methods for hyperbolic partial differential equations, in *Mathematical Aspects of Finite Elements in Partial Differential Equations: Symposium 1974*, ed. C. de Boor, Academic Press, 1974, pp. 195–212.
13. P. Lax and B. Wendroff, Systems of conservation laws, *Comm. Pure Appl. Math.* **13** (1960), 217–237.
14. R. I. McLachlan, Symplectic integration of Hamiltonian wave equations, *Numer. Math.* **66** (1994), 465–492.
15. 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.
16. R. I. McLachlan, G. R. W. Quispel, and N. Robidoux, Geometric integration using discrete gradients, *Phil. Trans. Roy. Soc. A,* **357** (1999), 1021–1046.
17. R. I. McLachlan and N. Robidoux, Antisymmetry, pseudospectral methods, and conservative PDEs, in *Proc. Int. Conf. EQUADIFF 99* pp. 994–999, B. Fiedler, K. Gröger, and J. Sprekels, eds., World Scientific, Singapore, 2000.
18. P. Olsson, Summation by parts, projections, and stability. I, *Math. Comput.* **64** (1995), 1035–1065, S23–S26.
19. G. R. W. Quispel and H. W. Capel, Solving ODE's numerically while preserving a first integral, *Phys. Lett. A* **218** (1996), 223–228.

20. R. Salmon and L. D. Talley, Generalizations of Arakawa's Jacobian, *J. Comput. Phys.* **83** (1989) 247–259.

21. C. V. Turner and R. R. Rosales, The small dispersion limit for a nonlinear semidiscrete system of equations, *Stud. Appl. Math.* **99** (1997), 205–254.