ANTISYMMETRY, PSEUDOSPECTRAL METHODS, WEIGHTED RESIDUAL DISCRETIZATIONS, AND ENERGY CONSERVING PARTIAL DIFFERENTIAL EQUATIONS

February 23, 2000 ROBERT I. MCLACHLAN AND NICOLAS ROBIDOUX

ABSTRACT. In view of its influence on the stability of numerical integrators, the *dual composition* method for producing energy-preserving discretizations of conservative PDEs, which generalizes the summation by parts approach, is introduced. Dual composition discretizations involve two compatible weighted residual approximations: one which targets variational derivatives of (nonlinear) functionals, and one which targets linear differential operators. In the simplest cases, it coincides with a pseudospectral, Galerkin, discretization. In general, however, the resulting differentiation matrices are not the standard pseudospectral—or the standard Galerkin—ones. Nonetheless, fast implementations can sometimes be recovered.

The emphasis is on pseudospectral bases; however, spectral and finite element bases are also discussed. Fourier antialiasing is shown to make some discretizations conservative; a Chebyshev grid analog is presented. Semi-discretizations of the one-way and the inhomogeneous wave equations on the interval are discussed. Several semi-discretizations, some of them Hamiltonian, of the Korteweg-de Vries equation on the circle are compared. The shortcomings of quadrature approximations of energy functionals are discussed.

CONTENTS

1. Introduction

For all $u, v \in C^1([-1,1])$,

$$\int_{-1}^{1} u\partial_x v\, dx = -\int_{-1}^{1} v\partial_x u\, dx + [uv]_{-1}^{1}, \tag{1}$$

so the operator ∂_x is skew-adjoint on $\{w \in C^1([-1,1]) : w(\pm 1) = 0\}$ with respect to the L^2 inner product $\langle\, ,\rangle$.

Take n points $x_i \in \mathbb{R}$, a real function $v(x)$, and estimate $v'(x_i)$ from the point values $v_i := v(x_i)$. In vector notation, $\mathbf{v}' = D\mathbf{v}$, where D is a "differentiation matrix." For example, with the grid of Chebyshev extreme points $x_i = -\cos(\frac{i\pi}{n-1})$, $i = 0, \ldots, n-1$, the Chebyshev pseudospectral differentiation matrix for $n = 4$—so that $(x_0, x_1, x_2, x_3) = (-1, -1/2, 1/2, 1)$—is

$$D = \frac{1}{6}\begin{pmatrix} -19 & 24 & -8 & 3 \\ -6 & 2 & 6 & -2 \\ 2 & -6 & -2 & 6 \\ -3 & 8 & -24 & 19 \end{pmatrix}, \tag{2}$$

which is not antisymmetric. This is typical: Chebyshev differentiation matrices are not antisymmetric (for all $n > 2$), unlike the Fourier spectral and centered differences differentiation matrices on uniform periodic grids. This has been observed many times in the literature, usually without comment other than it being unfortunate [3, 11, 13, 14, 18, 34, 35, 36, 37, 41]. This "failure" to be antisymmetric is puzzling in view of the (formal) skew-adjointness of ∂_x. Is the global nature of Chebyshev interpolation the problem? Is the lack of translation symmetry of the grid at fault? Is some key property of trigonometric functions missed by polynomials?[1] Perhaps the boundary values are at fault? If we restrict to functions

[1]It can be shown that a differentiation matrix obtained by interpolating, differentiating the interpolant, then evaluating the result at the grid points—as is done in pseudospectral and many finite difference methods—is antisymmetric if the interpolation operator is reflection invariant (if piecewise interpolants use local stencils, "local" reflection invariance is sufficient). Unless the interpolation stencil is very small, such interpolation operators are only practical on uniform

satisfying $w(\pm 1) = 0$, then the boundary terms in (1) vanish and ∂_x is skew-adjoint with respect to the L^2 inner product. The corresponding differentiation matrix is the central 2×2 block in (2), which is not antisymmetric either.

When discretizing conservative PDEs, antisymmetry matters for several reasons:

(i) Antisymmetry is necessary and sufficient for the construction of schemes with conserved quantities such as energy. For, if $\dot{\mathbf{v}} = D\nabla H(\mathbf{v})$ and D is antisymmetric, then
$$\dot{H} = (\nabla H)^{\mathrm{T}}\, \dot{\mathbf{v}} = (\nabla H)^{\mathrm{T}}\, D \nabla H = 0,$$
and conversely, if the flow of $\dot{\mathbf{v}} = \mathbf{f}$ preserves $H(\mathbf{v})$, then $\mathbf{f}^{\mathrm{T}}\nabla H \equiv 0$, and consequently
$$\mathbf{f} = \left\{ \frac{1}{|\nabla H|^2} \left(\mathbf{f}\,(\nabla H)^{\mathrm{T}} - (\nabla H)\,\mathbf{f}^{\mathrm{T}} \right) \right\} \nabla H,$$
the bracketed matrix being antisymmetric. (See [19, 20, 21, 22] for details and generalizations.)

(ii) Lack of antisymmetry leads to nonnormal matrices, which are not well characterized by their eigenvalues. Instead of a fairly straightforward eigenvalue analysis [14, 18, 31], the stability analysis of schemes based on nonnormal matrices involves more sophisticated techniques based on the Kreiss matrix theorem [14, 16, 18, 28, 31, 30, 38, 39] or pseudospectra [32, 33, 36].

(iii) Skew-adjointness is a structural feature associated with the differential operator ∂_x. Such features should be preserved under discretization if we want schemes that are natural and robust in the face of changes in the grid, the solution, or the operator.

The solution to the "puzzle" is that the inner product used in (1)—the L^2 inner product $\langle v, w \rangle := \int_{-1}^{1} vw\, dx$—must be discretized along with the operator ∂_x. Thus, suppose that the discrete approximation of the inner product $\langle\, ,\rangle$ is
$$\langle \mathbf{v}, \mathbf{w} \rangle_S \;:=\; \mathbf{v}^{\mathrm{T}} S \mathbf{w} \;\approx\; \int_{-1}^{1} vw\, dx \;=\; \langle\, , \rangle,$$
where S is symmetric positive definite. Then, for differentiation matrices of the form
$$D = S^{-1} A, \tag{3}$$
where $A = -A^{\mathrm{T}}$ is antisymmetric,
$$\langle \mathbf{v}, D\mathbf{w} \rangle_S + \langle D\mathbf{v}, \mathbf{w} \rangle_S \;=\; \mathbf{v}^{\mathrm{T}} S S^{-1} A \mathbf{w} + \mathbf{v}^{\mathrm{T}} A^{\mathrm{T}} S^{-\mathrm{T}} S \mathbf{w} \;=\; 0. \tag{4}$$
That is, D is skew-adjoint with respect to the inner product $\langle\, ,\rangle_S$. The question now becomes: which differentiation matrices D have factorizations of the form (3)?

The Chebyshev pseudospectral D shown in Eq. (2) factors as
$$\begin{aligned} D &= S^{-1} A \\ &= \frac{1}{512} \begin{pmatrix} 4096 & -304 & 496 & -1024 \\ -304 & 811 & -259 & 496 \\ 496 & -259 & 811 & -304 \\ -1024 & 496 & -304 & 4096 \end{pmatrix} \frac{1}{270} \begin{pmatrix} -135 & 184 & -72 & 23 \\ -184 & 0 & 256 & -72 \\ 72 & -256 & 0 & 184 \\ -23 & 72 & -184 & 135 \end{pmatrix}. \end{aligned} \tag{5}$$

---

periodic grids. Note that trigonometric interpolation and a uniform grid on the circle do *not* guarantee an antisymmetric differentiation matrix; a counterexample is given in Section 2.4.3.

The matrix A is not quite antisymmetric; its first and last diagonal entries correspond to the boundary term in (1). Thus,

$$\langle \mathbf{v}, D\mathbf{w}\rangle_S = -\langle D\mathbf{v}, \mathbf{w}\rangle_S + [vw]_{-1}^{1}, \tag{6}$$

and D is skew-adjoint with respect to functions for which the boundary term vanishes, just as in the continuous case. The construction which leads to Eq. (5) is carried out in a general setting in Section 2.

We systematically distinguish properties which are invariant under arbitrary change of basis from those which depend on a particular choice. The symmetry of S and antisymmetry of A are intrinsic, i.e., basis independent, unlike that of D itself. In the case of Fourier pseudospectral differentiation, for example, two "coincidences" (explained in Section 2) imply that $S = I$ in the standard basis, and therefore that the corresponding $D = S^{-1}A$ is A itself and consequently antisymmetric.

Although such factorizations are very well known in the field of finite element methods, they are not in other contexts. They appear to have been introduced for finite difference methods by Kreiss and Scherer [17] (see [6, 7, 24, 29] for developments). Although the coincidence of spectral collocation and Galerkin discretizations when the inner product and collocation points are "matched" is well known [3, 14], the earliest reference, that we know, to the existence of factorizations of the type (3) for arbitrary grids and inner products to Carpenter and Gottlieb [5].

Summation by parts formulas like Eq. (4) have been used to establish the stability of discretizations of linear PDEs by control of $\int u^2\,dx$ (the energy method) [5, 6, 7, 14, 15, 18, 24, 27, 29]. In Sections 4 and 6, we generalize the "summation by parts" approach to include nonlinear equations, differential operators other that $\mathcal{D} = \partial_x$, and energy functionals other that $\int u^2\,dx$, and shed light on standard pseudospectral, Galerkin, and antialiased pseudospectral methods.

The Fast Fourier Transform-based approach which makes pseudospectral methods desirable can be applied to evaluate Galerkin discretizations exactly in some cases. In Section 4, we show that if $\mathcal{D}(u)$ depends linearly on the field variable u, as it does in many Eulerian fluid equations, then an antialiased pseudospectral scheme does evaluate the Galerkin approximation exactly (this also holds if $\mathcal{D}(u)$ is polynomial in u). We claim that this is the origin of antialiasing's famous ability to suppress nonlinear instabilities [3]. (See, also, [26], in which antialising of the solution and analogously projected boundary conditions are used to control the instabilities of Chebyshev collocation discretizations of symmetric hyperbolic PDEs with variable coefficients.) This interpretation of antialiasing allows us to formulate a polynomial analogue of classical Fourier antialiasing.

The main application of summation-by-parts formula like Eq. (4) is to energy properties of numerical integration methods. For stability reasons, if the rate of change of the energy of continuous solutions depends only on boundary fluxes then the same should be true of the discretization. If the PDE is conservative, we want to know how to make the discretization conservative, or at least control the energy flux.

Let $\mathcal{H}(u)$ be a continuum conserved quantity, which we call the *energy.* In Section 6, for PDEs of the form

$$\dot{u} = \mathcal{D}(u)\frac{\delta \mathcal{H}}{\delta u}, \tag{7}$$

we construct corresponding "linear-gradient" spatial discretizations, namely ODEs of the form

$$\dot{\mathbf{u}} = L(\mathbf{u})\nabla H(\mathbf{u}) \tag{8}$$

with appropriate discretizations of u, $\mathcal{D}$, $\mathcal{H}$, and $\delta/\delta u$.

For a PDE of the form (7), if $\mathcal{D}(u)$ is formally skew-adjoint, then $d\mathcal{H}/dt$ depends only on the total energy flux through the boundary, which is often of a fixed sign; if $\mathcal{D}(u)$ is skew-adjoint, then the energy flux through the boundary is zero, and $\mathcal{H}$ is an integral. Analogously, if (8) holds,

$$\dot{H} \;=\; (\nabla H)^{\mathrm{T}}\dot{\mathbf{u}} \;=\; (\nabla H)^{\mathrm{T}} L \nabla H \;=\; (\nabla H)^{\mathrm{T}}\frac{1}{2}(L+L^{\mathrm{T}})\nabla H,$$

so that H cannot increase if the symmetric part of L is negative definite, and H is an integral if L is antisymmetric. (Conversely, all systems with an integral can be written in "skew-gradient" form, i.e., as in Eq. (8) with L antisymmetric [21, 22].)

Only for very simple problems, such as the wave equation, does the standard Galerkin approximation lead to a linear-gradient discretization. In Section 6, we introduce a general discretization method, based on compatible weighted residual discretizations of the linear operator $\mathcal{D}(u)$ and the variational derivative $\delta\mathcal{H}/\delta u$, which produces linear-gradient spatial discretizations of PDEs of the form (7). We call it the "dual composition" method. (Note that although we do not use the language of mixed finite element methods to describe it, the dual composition method does fit in the mixed FEM framework.)

Briefly, we use two finite-dimensional function spaces $\mathfrak{F}_0$ and $\mathfrak{F}_1$, where $\mathfrak{F}_0$ is "natural" for $\mathcal{H}$, and $\mathfrak{F}_1$ is a "domain of interior skewness" for $\mathcal{D}$. For $u \in \mathfrak{F}_0$, we approximate $\frac{\delta\mathcal{H}}{\delta u}$ in $\mathfrak{F}_1$ with weights (test functions) in $\mathfrak{F}_0$. For $v \in \mathfrak{F}_1$, we approximate $\mathcal{D}v$ in $\mathfrak{F}_0$, with test functions in $\mathfrak{F}_1$. This leads to discretizations

$$(\mathcal{D})(\frac{\delta\mathcal{H}}{\delta u}) \approx (S^{-\mathrm{T}}A)(S^{-1}\nabla H)$$

of linear-gradient form. Moreover, if $\mathcal{D}$ is skew-adjoint, then $A^{\mathrm{T}} = -A$, and the discretization is of skew-gradient form so that H is an integral. In Section 7, the new methods are applied to the wave equation, the Korteweg-de Vries (KdV) equation, and an inhomogeneous wave equation.

Hamiltonian systems are naturally in the form (7), with $\mathcal{H}$ the Hamiltonian and $\mathcal{D}(u)$ the Poisson operator. If $\mathcal{D}$ is constant—that is, independent of u—then so is A, and the discretization (8), with $L = S^{-\mathrm{T}}AS^{-1}$, is a Hamiltonian system. An earlier study [20] revealed that spectral, pseudospectral, and finite difference approximations on constant grids are Hamiltonian for constant Poisson brackets. However, only constant grids were considered, so the problems that arise on other grids and function spaces were not appreciated.

The theory applies to spectral, pseudospectral, and finite element methods with arbitrary grids and Hamiltonians. The energy $\mathcal{H}$ is approximated by $H(\mathbf{u}) := \mathcal{H}(u_j f_j)$, where $\{f_j\}$ is a basis of $\mathfrak{F}_0$. If $\mathcal{H}$ is complicated, this may be very expensive to compute, so we consider in Section 8 whether the nonlinear terms in $\mathcal{H}$ can be approximated by quadrature. In general, unfortunately, quadrature destroys the energy conserving structure of the discretization.

We do obtain conservative discretizations in four cases:

(i) when the energy $\mathcal{H}(u_j f_j)$ can be evaluated cheaply (say, in closed form) and $\mathcal{D}$ is skew-adjoint on $\mathfrak{F}_1$ (in some cases, cheap evaluation of $\frac{\delta\mathcal{H}}{\delta u}$ rather than $\mathcal{H}$ itself is sufficient);
(ii) when the grid is periodic and constant-spaced;
(iii) when the energy integral is approximated by quadrature and the basis $\{f_j\}$ is orthogonal, e.g., the cardinal functions on Legendre points on $[-1,1]$;
(iv) when the energy integral is approximated by quadrature and D commutes with the diagonal matrix of quadrature weights (for example, for Hamiltonian PDEs in canonical variables).

2. Differentiation matrices

Our primary aim is correct energy flux in the interior of the spatial domain. In general, there may be boundary contributions as in Eq. (1). This motivates the following.

2.1. Domains of skewness.

Definition 1. *Let $(\mathfrak{F},\langle\,,\rangle)$ be an inner product function space, $\mathfrak{F}_1 \subseteq \mathfrak{F}$ be a linear subspace, and $\mathcal{D}:\mathfrak{F}_1\to\mathfrak{F}$ be a linear operator.*

We say that $\mathcal{D}$ is formally skew-adjoint *(with respect to $\langle\,,\rangle$) if there is a functional $b(v,w)$, depending only on the values at the boundary of v and w and finitely many of their derivatives, such that*

$$\langle v,\mathcal{D}w\rangle = -\langle \mathcal{D}v,w\rangle + b(v,w),\ \forall v,w\in\mathfrak{F}_1.$$

$\mathfrak{F}_1$ is called a domain of interior skewness *of $\mathcal{D}$. If $b(v,w)=0$ for all v, $w\in\mathfrak{F}_1$, then $\mathfrak{F}_1$ is called a* domain of skewness *of $\mathcal{D}$, and we say that $\mathcal{D}$ is* skew adjoint.

If $\mathfrak{F}$ is omitted, we assume it is an appropriate L_2 space. If b is omitted, we assume that such a b exists. If the spatial domain has no boundary, then interior skewness coincides with skewness.

Definition 2. *The bilinear operator $\mathcal{A}:\mathfrak{F}_1\times\mathfrak{F}_1\to\mathbb{R}$ is defined by*

$$\mathcal{A}(v,w) := \langle v,\mathcal{D}w\rangle\,. \tag{9}$$

Proposition 1. *$\mathcal{A}$ is antisymmetric if and only if $\mathcal{D}$ is skew-adjoint with respect to $\langle\,,\rangle$.*

We illustrate the key ideas with the differential operator $\mathcal{D}=\partial_x$.

Example 1 (Domains of skewness of $\mathcal{D}=\partial_x$). Let $\mathfrak{C} = C^1([-1,1])$. Eq. (1) shows that the following are domains of interior skewness of $\mathcal{D}=\partial_x$, with (i)–(iv) being fully fledged domains of skewness. Examples (i)–(v) are infinite dimensional function spaces, while (vi)–(viii) are finite dimensional.

(i) $\mathfrak{F}_1 = C^1(\mathbb{S}^1) = \{v\in\mathfrak{C}: v(1)=v(-1),\ v_x(1)=v_x(-1)\}$ ("periodic").
(ii) $\mathfrak{F}_1 = \{v\in\mathfrak{C}: v(\pm1)=0\}$ ("zero boundary values").

There is no domain of skewness containing both of the following domains of skewness. Thus, in general, there is no unique maximal domain of skewness.

(iii) $\mathfrak{F}_1 = \{v\in\mathfrak{C}: v(1)=v(-1)\}$ ("equal boundary values").
(iv) $\mathfrak{F}_1 = \{v\in\mathfrak{C}: v(1)=-v(-1)\}$ ("opposite boundary values").

Although $\mathcal{D}$ is the differentiation operator ∂_x, functions in the domain of skewness need not be differentiable everywhere.

(v) Let $-1 = x_0 < x_1 < \cdots < x_n = 1$ be a partition of $[-1,1]$ with interior $P = [-1,1] \setminus \cup\{x_i\}$.
$\mathfrak{F}_1 = C^1(P) \cap C^0([-1,1])$ ("piecewise C^1"), $b(v,w) = [vw]_{-1}^1$.

The next three examples are finite-dimensional function spaces used in numerical approximations: $\mathfrak{F}_1^{\rm trig}$ in spectral methods, $\mathfrak{F}_1^{\rm poly}$ in pseudospectral methods, and $\mathfrak{F}_1^{\rm pp}$ in finite element methods.

(vi) $\mathfrak{F}_1^{\rm trig}(n) = \{v : v = a_0 + \sum_{k=1}^n a_k \sin(kx) + b_k \cos(kx)\}$ ("trigonometric polynomials of degree n"), $b(v,w) = [vw]_{-1}^1$.

(vii) $\mathfrak{F}_1^{\rm poly}(n) = \mathfrak{P}_n([-1,1])$. ("polynomials of degree n"), $b(v,w) = [vw]_{-1}^1$.

(viii) $\mathfrak{F}_1^{\rm pp}(n,r) = \{v \in C^r([-1,1]) : v|_{[x_i,x_{i+1}]} \in \mathfrak{P}_n([x_i,x_{i+1}])\}$ ("piecewise polynomials of degree n with r derivatives"), $b(v,w) = [vw]_{-1}^1$.

In general, $\mathcal{D} = \partial_x$ does not map $\mathfrak{F}_1$ into itself. Examples (vi) and (vii), $\mathfrak{F}_1^{\rm trig}$ and $\mathfrak{F}_1^{\rm poly}$, are the only ones with $\mathcal{D}(\mathfrak{F}_1) \subseteq \mathfrak{F}_1$.

For numerical work, a basis of $\mathfrak{F}_1$ is chosen. We write $\{f_j\}$ for a basis of $\mathfrak{F}_1$. Given a grid $\{x_i\}$, a *cardinal basis* is one for which $f_j(x_i) = \delta_{ij}$. For $\mathfrak{P}_n([-1,1])$, useful bases are the Legendre polynomials, the Chebyshev polynomials of the first kind $\{T_j(x) : 0 \le j \le n\}$, and, given a grid, the cardinal basis of Lagrange interpolating polynomials. If $\langle f_i, f_j\rangle = \delta_{ij}$ for all i and j, the basis is said to be *orthonormal* (with respect to $\langle , \rangle$).

Any $v \in \mathfrak{F}_1$ can be expanded uniquely as

$$v(x) = \beta_j f_j(x)$$

(we sum on repeated indices). For a cardinal basis, $\beta_j = v(x_j) := v_j$. We collect the ("spectral") coefficients β_j into a column vector $\boldsymbol{\beta}$, and the point values v_j into a column vector $\mathbf{v}$.

Definition 3. *Let $\{f_j\}$ be a basis of $\mathfrak{F}_1$. $L_{ij} := \mathcal{L}(f_i, f_j)$ is called the (i,j)* matrix element *of the bilinear operator $\mathcal{L} : \mathfrak{F}_1 \times \mathfrak{F}_1 \to \mathbb{R}$. If $\mathfrak{F}_1$ has finite dimension n, then $L = (L_{ij}) \in \mathbb{R}^{n\times n}$ is the* matrix *of $\mathcal{L}$ with respect to the basis.*

Definition 4. *S is the matrix of the inner product $\langle , \rangle$.*

Proposition 2. *S is symmetric positive definite. The basis is orthonormal if and only if $S = I$, the identity matrix.*

Definition 5. *A is the matrix of the bilinear operator $\mathcal{A}$.*

Proposition 3. *$\mathfrak{F}_1$ is a finite dimensional domain of skewness of $\mathcal{D}$ if and only if A is antisymmetric.*

2.2. Differentiation in finite dimensional function spaces.

Two things cause differentiation in the space $\mathfrak{F}_1$ to be nontrivial. Firstly, we may not have $\mathcal{D}(\mathfrak{F}_1) \subseteq \mathfrak{F}_1$. Secondly, functions in $\mathcal{D}(\mathfrak{F}_1)$ may not satisfy the same boundary conditions as functions in $\mathfrak{F}_1$.

The only reasonably general method for constructing approximations that respect the symmetry of the underlying operators was found by the authors to be closely related to the weighted residual method [3, 10, 42], a generalization of the well-known Galerkin finite element method [3, 10, 14, 42].

Definition 6. *Let $\mathfrak{F}_1$ be an inner product function space. The Galerkin approximation of $\mathcal{D}v$ is the unique function v' such that $\langle w, v'\rangle = \mathcal{A}(w,v)$ for all $w \in \mathfrak{F}_1$.*

We require the residual $v' - \mathcal{D}v$ to be orthogonal to $\mathfrak{F}_1$, *with respect to the inner product which makes $\mathcal{D}$ (formally) skew-adjoint.* For accuracy, it is important that $\mathfrak{F}_1$ be a reasonable space in which to approximate $\mathcal{D}v$ for $v \in \mathfrak{F}_1$.

Compare this to another standard spectral method in which the residual is required to be orthogonal with respect to the inner product in which the basis is orthogonal. The Chebyshev polynomials, for example, are orthogonal with respect to the inner product $\langle u, v\rangle = \int_{-1}^{1} uv(1-x^2)^{-1/2}\,dx$, *not* the L^2 inner product. Such methods are convenient since they lead to explicit approximations because S is then diagonal and consequently S^{-1} trivial to compute. Typically, however, they do not lead to symmetry-preserving methods, because the inner product was chosen to suit the basis, not the operator $\mathcal{D}$.

For a Galerkin method, the test functions are equal to the trial functions. Expanding in a basis $\{f_j\}$ of $\mathfrak{F}_1$, let $v = \beta_j f_j$, $v' = \beta_j' f_j$, and $w = \gamma_j f_j$. Then

$$\boldsymbol{\gamma}^{\mathrm{T}} S \boldsymbol{\beta}' = \boldsymbol{\gamma}^{\mathrm{T}} A \boldsymbol{\beta}, \ \forall \boldsymbol{\gamma}.$$

Since S is nonsingular,

$$\boldsymbol{\beta}' = S^{-1} A \boldsymbol{\beta}.$$

Note that in practice it is not always necessary to compute S and A: one can instead compute $\mathcal{D}v$ exactly and project it orthogonally onto $\mathfrak{F}_1$.

Definition 7. *The* Galerkin differentiation matrix D *associated with the inner product* $\langle\, ,\rangle$ *and basis* $\{f_j\}$ *is* $D := S^{-1}A$.

When the functions f_j are nonlocal, D is often called the spectral differentiation matrix. The link to standard pseudospectral methods is that some Galerkin methods are pseudospectral.

Proposition 4.
(i) *If $\mathcal{D}(\mathfrak{F}_1) \subseteq \mathfrak{F}_1$, then $v' = \mathcal{D}v$, i.e., the Galerkin approximation of the derivative is exact.*
(ii) *If $\mathcal{D}(\mathfrak{F}_1) \subseteq \mathfrak{F}_1$ and $\{f_j\}$ is a cardinal basis, then D is the standard pseudospectral differentiation matrix.*

Proof.
(i) We have $\langle w, v' - \mathcal{D}v\rangle = 0$ for all $w \in \mathfrak{F}_1$. If $\mathcal{D}v \in \mathfrak{F}_1$, then $v' = \mathcal{D}v$ is a solution. But solutions are unique because S is nonsingular.
(ii) From (i), the derivative is exact. We have $v = \beta_j f_j$, so $\mathcal{D}v = \beta_j \mathcal{D} f_j$. In a cardinal basis, spectral coefficients and point values coincide, so the spectral coefficients of the derivative are $\mathcal{D}f_j(x_i)\beta_j$, i.e., $D_{ij} = \mathcal{D}f_j(x_i)$. This is exactly the standard pseudospectral differentiation matrix. ☐

We emphasize the following points—well known in the finite element context—in the pseudospectral context:
(i) Galerkin differentiation depends only on $\mathfrak{F}_1$, i.e., it is basis independent.
(ii) The matrices A, S, and D depend on the chosen basis, but the factorization $D = S^{-1}A$ is true in any basis.
(iii) The construction is independent of the grid.
(iv) The factorization $D = S^{-1}A$ allows us to study antisymmetry.

Proposition 5. *If $\mathfrak{F}_1$ is a domain of skewness, then*
(i) *D is skew-adjoint with respect to the inner product induced by S,*

(ii) *D is similar to an antisymmetric matrix, and*
(iii) *D is antisymmetric if the basis is orthonormal.*

Proof. By Proposition 3, A is antisymmetric. Let $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ be column vectors.

(i)
$$\langle \boldsymbol{\beta}, D\boldsymbol{\gamma}\rangle_S \;=\; \boldsymbol{\beta}^{\mathrm{T}} S S^{-1} A \boldsymbol{\gamma} \;=\; \boldsymbol{\beta}^{\mathrm{T}} A \boldsymbol{\gamma} \;=\; -\boldsymbol{\gamma}^{\mathrm{T}} A \boldsymbol{\beta} \;=\; -\langle D\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle_S .$$

(ii) Since S is symmetric positive definite, it has a unique symmetric positive definite square root $S^{\frac{1}{2}}$. Therefore
$$D \;=\; S^{-1}A \;=\; S^{-\frac{1}{2}}\left(S^{-\frac{1}{2}} A S^{-\frac{1}{2}}\right) S^{\frac{1}{2}}.$$

(iii) If the basis is orthonormal, $S = I$, so $D = A$, which is antisymmetric if $\mathfrak{F}_1$ is a domain of skewness. □

The matrix A is not always antisymmetric, because of the boundary terms. However, it is antisymmetric up to boundary terms, in the following sense.

Proposition 6. *Let $\mathfrak{F}_1$ be a domain of interior skewness of $\mathcal{D}$. Then $\mathcal{A}$ and A are antisymmetric when evaluated on functions satisfying skew boundary conditions, that is, $\mathcal{A}(v,w) = -\mathcal{A}(w,v)$ and $\boldsymbol{\beta}^{\mathrm{T}} A \boldsymbol{\gamma} = -\boldsymbol{\beta}^{\mathrm{T}} A^{\mathrm{T}} \boldsymbol{\gamma}$ for all $v = \beta_j f_j$, $w = \gamma_j f_j$ such that $b(v,w) = 0$.*

Proof. True for $\mathcal{A}$ by definition. A is the representation $\mathcal{A}$ with respect to the basis. □

2.3. **Change of basis.** The transformation rules for a multilinear operator (differentiation, inner product, etc.) depend on its type and determine its normal form and invariants. For example, the linear map $x \mapsto Lx$ transforms under $y = Cx$ to $y \mapsto CLC^{-1}y$; L is transformed to CLC^{-1}. The invariants associated with L are therefore its eigenvalues, and L can be transformed to its Jordan normal form.

Let $\{f_j\}$ and $\{\tilde{f}_i\}$ be two bases of the linear function space $\mathfrak{F}_1$. Let A (resp. $\widetilde{A}$) be the matrix of $\mathcal{A}$ in the basis $\{f_j\}$ (resp. $\{\tilde{f}_i\}$), etc. Each function $\tilde{f}_i$ can be expanded as
$$\tilde{f}_i = C_{ij} f_j.$$
Then,
$$\widetilde{S}_{ij} \;=\; \int \tilde{f}_i \tilde{f}_j \, dx \;=\; \int C_{ik} f_k C_{jl} f_l \, dx \;=\; C_{ik} S_{kl} C_{jl},$$
that is,
$$\widetilde{S} = CSC^{\mathrm{T}}. \tag{10}$$
Similarly,
$$\widetilde{A} = CAC^{\mathrm{T}}, \tag{11}$$
while
$$\widetilde{D} \;=\; \widetilde{S}^{-1}\widetilde{A} \;=\; C^{-\mathrm{T}} S^{-1} C^{-1} C A C^{\mathrm{T}} \;=\; C^{-\mathrm{T}} S^{-1} A C^{\mathrm{T}} \;=\; C^{-\mathrm{T}} D C^{\mathrm{T}}.$$
A and S transform as $(0,2)$ tensors, D as a $(1,1)$ tensor, that is, a linear map. If A and S are antisymmetric, so are $\widetilde{A}$ and $\widetilde{S}$, *regardless of C*. However, if D is antisymmetric, then $\widetilde{D}$ is antisymmetric if and only if D commutes with $C^{\mathrm{T}}C$, for

example, if C is orthogonal. Antisymmetry is an intrinsic property of $(0,2)$ tensors but not of linear maps.

The invariants of D are its eigenvalues. S has no invariants: there is a basis in which $S = I$. If A is antisymmetric, its only invariant is its rank: Darboux's theorem (see, e.g., [25]) states that it can be brought into the normal form

$$\begin{pmatrix} 0 & I & 0 \\ -I & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

for submatrices of appropriate size.

If the eigenvalues of A are not invariant under changes of variables, do they have any meaning at all? In a sense they do; for, since A is antisymmetric, it is normal, hence unitarily diagonalizable, say to $UAU^{-1} = \Lambda$. Since U is unitary, Λ is a valid transformation of A. The imaginary conjugate pairs of eigenvalues, which can be reduced to $\pm i$ by a (nonunitary) change of variables, reflect the nonorthonormality of successive pairs of conjugate variables.

Example 2 (**Orthonormalization**)**.** Let $\{\tilde{f}_i\}$ be any orthonormal basis, constructed by, e.g, Gram-Schmidt orthonormalization of $\{f_j\}$, or by $\tilde{f}_i = S_{ij}^{-\frac{1}{2}} f_j$. Then $\widetilde{S} = I$ and $S = C^{-1}C^{-\mathrm{T}}$. This can be viewed as a way to evaluate S, or as a way to construct an antisymmetric $\widetilde{D}$.

Example 3 (**Spectral coefficients to point values**)**.** Cardinal bases have the nice property that $\tilde{f}_i(x_k) = \delta_{ik}$, so that the expansion coefficients with respect to a cardinal basis correspond to point values. With a cardinal basis, if the boundary points are included in the grid, the boundary term $[vw]_{-1}^{1}$ in Eq. (6) equals $v_n w_n - v_0 w_0$.

Let $\{f_j\}$ be an arbitrary basis and $\{\tilde{f}_i\}$ be a cardinal basis for the grid points $\{x_i\}$. Then $\tilde{f}_i = C_{ij} f_j$ gives $\delta_{ik} = \tilde{f}_i(x_k) = C_{ij} f_j(x_k)$, i.e., $I = CF$, where $F_{jk} := f_j(x_k)$. Consequently, A and S can be calculated in a convenient basis, and then transformed to a cardinal basis using Eq. (10) and (11) with $C = F^{-1}$.

Example 4 (**Spectral weights**)**.** Instead of the spectral coefficients β_i, one might want to use the spectral weights

$$\tilde{\beta}_i := \int v f_i \, dx.$$

The new basis functions are $\tilde{f}_i = S_{ij}^{-1} f_j$. Thus, working in terms of weights is equivalent to a change of basis with $C = S^{-\mathrm{T}} = S^{-1}$. Only when the original basis is orthonormal are weights and coefficients actually equal. The differentiation matrix with respect to the spectral weights is AS^{-1}.

An alternate definition of the Galerkin differentiation matrix in terms of the spectral weights can be given: If $v = \beta_j f_j$, then $v' = \mathcal{D}v = \beta_j \mathcal{D} f_j$, and it is natural to set the spectral weights of each side equal, so that, as before, $\tilde{\boldsymbol{\beta}}' = A\boldsymbol{\beta} = AS^{-1}\tilde{\boldsymbol{\beta}}$.

The matrix factors S and A are, we believe, more fundamental than D itself in any study of pseudospectral differentiation. Similar factorizations exist on arbitrary grids (as shown in [5], where the Legendre grid is studied in some detail). This is important, because the accuracy of numerical differentiation, unlike that of quadrature, is not strongly dependent on the grid [11].

Proposition 7. *A square matrix D is skew-adjoint (that is, there exists a positive definite hermitian matrix S and a skew-hermitian matrix A such that $D = S^{-1}A$) if and only if D is diagonalizable and its spectrum is purely imaginary.*

Proof. Suppose that D is skew-adjoint. Because S is positive definite hermitian, it has a positive definite hermitian square root K. Consider

$$\widetilde{D} := K^{-1}AK^{-1}.$$

Note that $D = K^{-1}\widetilde{D}K$. Because diagonalizability and spectrum are invariant under similarity transformations, it suffices to consider $\widetilde{D}$. $\widetilde{D}$ is normal, and consequently diagonalizable, since

$$\begin{aligned}\widetilde{D}\widetilde{D}^* - \widetilde{D}^*\widetilde{D} &= K^{-1}AK^{-1}K^{-*}A^*K^{-*} - K^{-*}A^*K^{-*}K^{-1}AK^{-1} \\ &= -K^{-1}AS^{-1}AK^{-1} + K^{-1}AS^{-1}AK^{-1} \;=\; 0.\end{aligned}$$

$\widetilde{D}$ has purely imaginary spectrum because it is skew-hermitian.

Conversely, suppose that D is diagonalizable and its spectrum is purely imaginary, so that $D = KiLK^{-1}$ with L diagonal and real. Then,

$$D = (KK^*)(iK^{-*}LK^{-1}).$$

The first factor is hermitian positive definite, the second, skew-hermitian. □

Although the eigenvalues of A and S do not have "physical" meaning, they do matter when studying the stability of schemes based on a family of matrices of the form $D = S^{-1}A$ with the help of generalizations of the Kreiss matrix theorem as found in [14, 18, 31, 38, 39].

2.4. **Fourier differentiation.** We now examine the structure of spectral differentiation matrices in detail. Of course, their properties are already well known. Our purpose is to shed light on why the spectral differentiation matrices are antisymmetric. This is the simplest setting in which to study the above theory, because no boundaries "get in the way".

2.4.1. *Fourier differentiation matrices of odd dimensions.* Consider periodic functions on the circle $\mathbb{S}^1$, identified with the interval $[-\pi, \pi)$. $C^1(\mathbb{S}^1)$ is a domain of skewness for $\mathcal{D} = \partial_x$. A $(2n+1)$-dimensional subspace $\mathfrak{F}_1$ is spanned by

$$f_1 = \frac{1}{\sqrt{2\pi}} \text{ and } f_{2k} = \frac{1}{\sqrt{\pi}}\cos(kx),\ \ f_{2k+1} = \frac{1}{\sqrt{\pi}}\sin(kx),\ \ k = 1, \dots, n. \tag{12}$$

This basis is orthonormal. With respect to this trigonometric basis, $S = I$ and $D = A$ is antisymmetric (Propositions 3 and 5 part (iii)); specifically, A is block diagonal:

$$D \;=\; A \;=\; \operatorname{diag}\left(0, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & n \\ -n & 0 \end{pmatrix}\right). \tag{13}$$

Unlike the antisymmetry of D, the antisymmetry of A depends only on $\mathcal{D}$ and $\mathfrak{F}_1$, not on the choice of basis (Proposition 3). For example, if we make the single change $\hat{f}_2 = \gamma f_2$, then the corresponding spectral differentiation matrix

$$\widehat{D} = \operatorname{diag}\left(0, \begin{pmatrix} 0 & \gamma \\ -1/\gamma & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & n \\ -n & 0 \end{pmatrix}\right)$$

is no longer antisymmetric, but its factor $\widehat{A}$ still is—the only change is that $\widehat{A}_{23} = -\widehat{A}_{32} = \gamma$.

What about the Fourier *pseudo*spectral differentiation matrix? Select a set of grid points $\{x_i\}$. Since $\mathcal{D}(\mathfrak{F}_1) \subseteq \mathfrak{F}_1$, the pseudospectral differentiation matrix $\widetilde{D}$ is the Galerkin (= spectral) differentiation matrix expressed in the cardinal basis (Proposition 4 part (ii)). Transforming as in Example 3, we see that the classical pseudospectral differentiation matrix $\widetilde{D}$ has the form

$$\widetilde{D} = \widetilde{S}^{-1}\widetilde{A}, \tag{14}$$

where

$$\widetilde{S} = F^{-1}SF^{-\mathrm{T}} = F^{-1}F^{-\mathrm{T}}$$

and

$$\widetilde{A} = F^{-1}AF^{-\mathrm{T}}.$$

So,

$$\widetilde{D} = F^{\mathrm{T}}AF^{-\mathrm{T}}.$$

If the grid is uniform, the matrix F satisfies

$$FF^{\mathrm{T}} = \frac{n}{2}I$$

(it is the real form of the matrix of the complex basis e^{ijx}, which is circulant and for which FF^{T} can be evaluated using geometric series). Therefore,

$$\widetilde{S} = \frac{2}{n}I,$$

so that the cardinal basis functions are orthonormal up to a uniform scaling, S is a multiple of the identity, and $\widetilde{D}$ is antisymmetric.

2.4.2. *Fourier differentiation matrices of even dimensions.* The situation is analogous for the standard Fourier spectral and pseudospectral differentiation matrices of *even* dimensions—frequently used because the FFT can be performed very efficiently on a grid with a power of two number of points using the Cooley-Tukey algorigthm [8]—even though in this case $\mathcal{D}(\mathfrak{F}_1) \not\subseteq \mathfrak{F}_1$. In the spectral case—using the trigonometric basis (12) minus the last sine mode—D is still equal to A which is antisymmetric because the basis is still orthonormal and its span a domain of skewness; A is as in Eq. (13) except that the last diagonal block is a 1×1 zero matrix. In the pseudospectral case, things work out as above because the derivative of $\cos(nx)$ vanishes at all points of the customary uniform grid (the zeros of $\sin(nx)$): Eq. (14) still holds, and $\widetilde{D}$, again, is antisymmetric.

2.4.3. *Trigonometric interpolation and a uniform grid do not guarantee an antisymmetric differentiation matrix.* Unlike spectral differentiation matrices, pseudospectral differentiation matrices are invariant with respect to rescalings of the basis functions (which are used for interpolation). If the number of grid points is odd, the Fourier pseudospectral differentiation matrix is antisymmetric no matter which uniform grid is used—that is, even if the grid is not "anchored" at 0—because the corresponding interpolation operator commutes with reflections and ∂_x anticommutes.

However, even in this simplest of situation—uniform periodic grid and trigonometric basis functions—one can see how fragile the antisymmetry of the pseudospectral differentiation matrix is. With the basis (12) truncated at an even number of basis functions, we have $\mathcal{D}(\mathfrak{F}_1) \not\subseteq \mathfrak{F}_1$; consequently, Proposition 4 cannot be used to show that the pseudospectral differentiation matrix is identical to the Galerkin one expressed in the cardinal basis; indeed, in Section 2.4.2, we had to call on the fact that the derivative of $\cos(nx)$ vanishes at all points of the customary uniform grid. This, however, does not hold in general: generically, the cardinal basis Galerkin and pseudospectral discretizations are distinct, and the pseudospectral differentiation matrix is not antisymmetric.

For example, keeping the same basis functions for interpolation as in Section 2.4.2, but translating the usual uniform grid by θ, so that

$$x_i = \left(-1 + \frac{i-1}{n}\right)\pi + \theta,\ i = 1, \ldots, 2n,$$

produces a non-antisymmetric pseudospectral differentiation matrix $\widetilde{D}$ for almost all θ's. To wit, let

$$v_i = \cos(n\theta)\sin(nx_i),\ i = 1, \ldots, 2n. \tag{15}$$

We do not allow $\cos(n\theta)$ to vanish because otherwise the interpolation problem does not always have a solution in the span of the basis functions. Suppose also that $\sin(n\theta) \neq 0$, so that $\mathbf{v} \neq \mathbf{0}$. We show that $\mathbf{v}$ is a real eigenvector of $\widetilde{D}$ with non-zero (real) eigenvalue, so that $\widetilde{D}$ cannot be antisymmetric, and actually cannot be skew-adjoint with respect to any inner product.

Because

$$\sin(nx)\cos(n\theta) - \cos(nx)\sin(n\theta) = \sin(n(x-\theta))$$

vanishes at all the grid points, the function in the span of the basis functions which interpolates the point data (15) is $\sin(n\theta)\cos(nx)$. Consequently, $\mathbf{v}' := \widetilde{D}\mathbf{v}$ is given by

$$v_i' \;=\; \partial_x(\sin(n\theta)\cos(nx))|_{x=x_i} \;=\; -n\sin(n\theta)\sin(nx_i),$$

and

$$\widetilde{D}\mathbf{v} = -n\tan(n\theta)\mathbf{v}.$$

For example, the matrix of the Fourier pseudospectral discretization of ∂_x using the grid $\{-\pi+\theta, \theta\}$ and interpolation basis $\{1, \cos(x)\}$ is

$$\widetilde{D} = \frac{1}{2}\tan(\theta)\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},$$

which is antisymmetric only if $\sin(\theta) = 0$.

2.5. **Polynomial differentiation.** Differentiation on spaces of polynomials on $[-1,1]$ is very similar to Fourier differentiation. The twist is the presence of boundary conditions in the domains of skewness.

The theory guarantees the factorization $D = S^{-1}A$ for any basis. $\mathfrak{F}_1 = \mathfrak{P}_n([-1,1])$ is a domain of interior skewness which is closed under $\mathcal{D} = \partial_x$. Seeking a form which manifests the influence of the boundary, we use a cardinal basis over a grid with points at the boundary.

The symmetric part of the matrix A in the factorization (5) of the Chebyshev differentiation matrix—the two elements $\pm 135/270 = \pm 1/2$—is shown in the following proposition to be solely a consequence of having grid points at the boundary (given that the basis functions are smooth).

Proposition 8. *Let $\{f_j\}$ be the Lagrange polynomials on the grid $-1 = x_0 < x_1 < \cdots < x_n = 1$, that is, $f_j(x_i) = \delta_{ij}$. Then the symmetric part of A,* $\mathrm{sym}(A) := (A + A^{\mathrm{T}})/2$, *is zero except for* $\mathrm{sym}(A)_{nn} = -\mathrm{sym}(A)_{00} = 1/2$.

Proof.

$$\mathrm{sym}(A)_{ij} = \frac{1}{2}\int f_i \mathcal{D} f_j + f_j \mathcal{D} f_i \, dx = \frac{1}{2}[f_i f_j]_{-1}^{1} = \begin{cases} -\frac{1}{2}, & i = j = 0, \\ \frac{1}{2}, & i = j = n, \\ 0, & \text{otherwise.} \end{cases} \qquad \square$$

Using the Chebyshev extreme points $x_i = \cos(i\pi/n)$, $i = 0, \ldots, n$—listed, this time, from right to left, as is usual when using cosine transforms—A can be evaluated in closed form with respect to the non cardinal basis $\{T_i\}$ of Chebyshev polynomials, and then transformed to the cardinal basis. We have

$$S_{ij}^{\mathrm{cheb}} = \int_{-1}^{1} T_i T_j \, dx = \begin{cases} 0, & i-j \text{ odd} \\ \dfrac{-2(i^2 + j^2 - 1)}{((i+j)^2 - 1)((i-j)^2 - 1)}, & i-j \text{ even} \end{cases}$$

[26] and

$$A_{ij}^{\mathrm{cheb}} = \int_{-1}^{1} T_i \mathcal{D} T_j \, dx = \begin{cases} \dfrac{2j^2}{j^2 - i^2}, & i-j \text{ odd} \\ 0, & i-j \text{ even} \end{cases}$$

with change of basis $F_{ij} = T_j(x_i) = \cos(ij\pi/n)$, a discrete cosine transform; then $A = F^{-1} A^{\mathrm{cheb}} F^{-\mathrm{T}} = F^{-1} A^{\mathrm{cheb}} F^{-1}$. ***NICOLAS: PUT IN THE CLOSE FORMS.***

Figure 1 shows the matrices D, A, S and S^{-1} for $n = 15$. S and A are very regular, both on and off the diagonal, and show less "wild behavior" than D, which is dominated by huge entries near the boundary points ± 1.

Because D is exact on $\mathfrak{P}_n$, and the $n+1$st derivative of polynomials of order $n+1$ vanishes, D is nilpotent, and its only eigenvalue is zero with geometric multiplicity 1, independently of the chosen basis of $\mathfrak{P}_n$. Consequently, the Jordan block decomposition of D is 0's everywhere but on the first upper diagonal, and similarity transformations which produces the Jordan form are readily available: given *any* polynomial p of degree (exactly) n, one can take the jth column of the transformation matrix to be the point values of the $n-j+1$st derivative of p (this, again, holds irrespective of the grid, and, replacing "point values" by "expansion coefficients," for any basis of $\mathfrak{P}_n$).

D is not diagonalizable since its Jordan block decomposition is not diagonal. Consequently, there is no matrix which induces an inner product with respect to which D is skew-adjoint by Proposition 7.

Since A is a real matrix, its spectrum is symmetric about the real axis; using that ∂_x anticommutes with reflections and that the grid is symmetric about the origin, one can show that A's spectrum is symmetric about the origin, and consequently the imaginary axis, as well. Because the nullspace of A is the same as the nullspace of D, the geometric multiplicity of the eigenvalue 0—corresponding to the Casimir of ∂_x—is always 1; the algebraic multiplicity of 0, on the other hand, appears to be 1 when A's dimensions are odd, 2 when even. The other eigenvalues of A appear to be single, purely imaginary, and fill $(-i\pi, i\pi)$ uniformly as $n \to \infty$ (Figure 2); for example, the largest two eigenvalues of A, for $n = 1023$, are $\pm 3.13i$. It appears

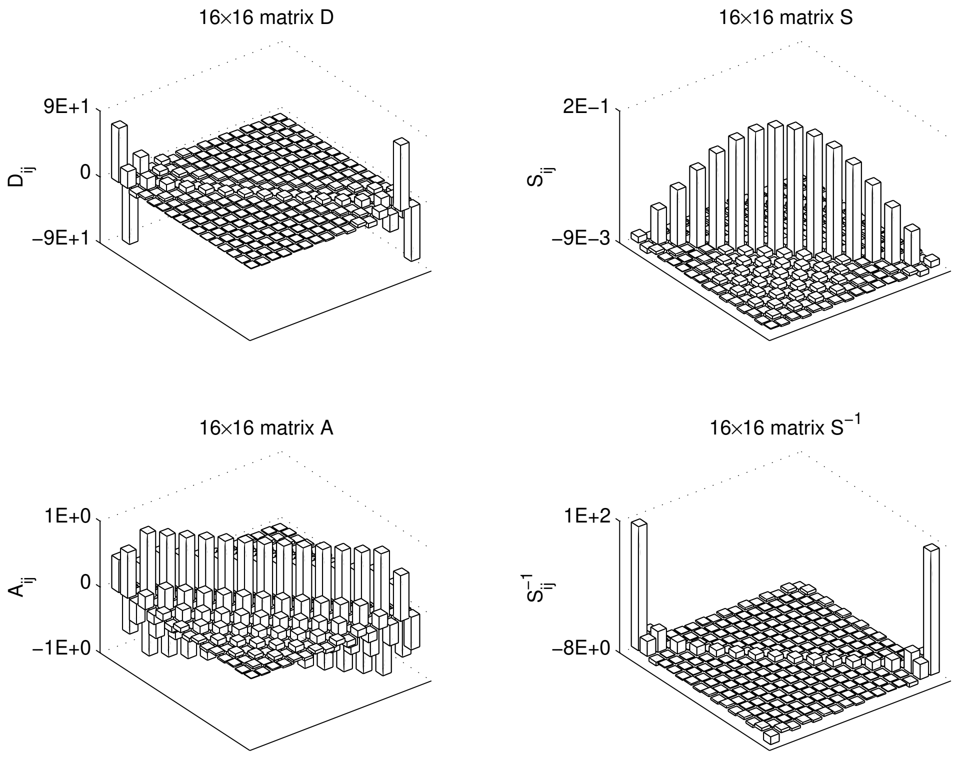

Figure 1: Matrix entries of the Chebyshev differentiation matrix D and its factors S, A and S^{-1} ($D = S^{-1}A$), for $n = 15$.

that the non-zero part of the spectrum of A is $\{i\lambda_1, -i\lambda_1, \ldots, i\lambda_m, -i\lambda_m\}$ with

$$\lambda_j \approx \frac{2\pi}{n+3} j \tag{16}$$

for n odd (A's with even dimensions), and

$$\lambda_j \approx \frac{2\pi}{n+3}\left(j - \frac{1}{2}\right) \tag{17}$$

for n even. For $n = 1023$ for example, the ratio between λ_j and $2\pi j/(n+3)$ which deviates the most from 1 is 1.0003.

The eigenvalues of S very closely approximate the quadrature weights of the Chebyshev grid (Figure 3). The ratios of the four smallest eigenvalues and the four smallest quadrature weights, for $n = 63$, are 1.85, 1.89, 1.33 and 1.36: even the extreme eigenvalues show little "spurious" behaviour. This is not too surprising, given that S and the diagonal matrix of quadrature weights have the same row and column sums, and that, as approximations of the L^2 inner product, they are exact on $\mathfrak{P}_m \times \mathfrak{P}_m$, where m is the largest integer such that $m \leq n/2$.

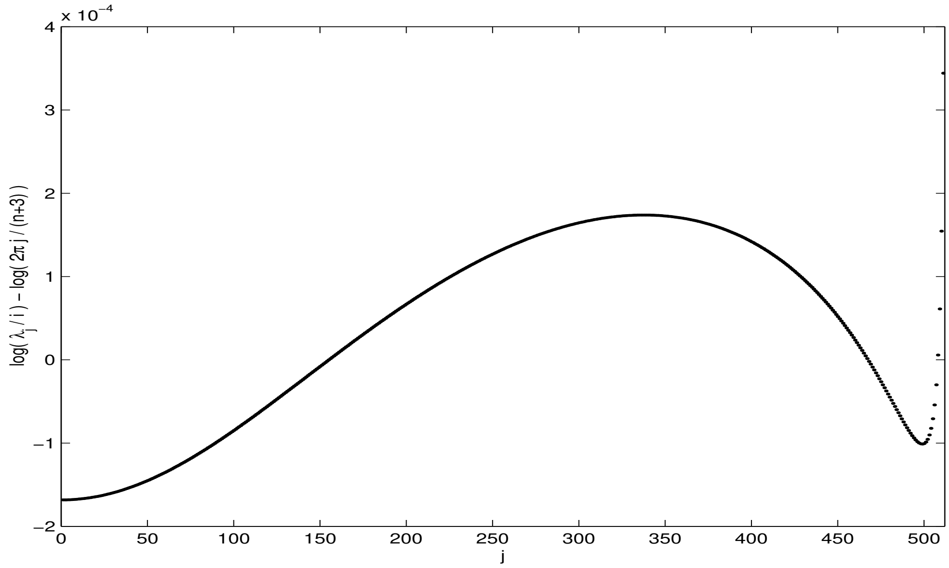

Figure 2: Logarithms of the ratios of the non-zero eigenvalues $\pm i\lambda_j$ of the Chebyshev differentiation matrix factor A, and of $2\pi ij/(n+3)$, for $n = 1023$ (non-zero eigenvalues, which appear to be single and purely imaginary, come in conjugate pairs). All the ratios are very close to 1, which supports Eq. (16). The corresponding graph for $n = 1022$, using Eq. (17) instead of Eq. (16), is a carbon copy.

2.6. **Finite elements.** For piecewise polynomial functions $\mathfrak{F}^{\rm pp}(n,r)$, spectral differentiation boils down to the Galerkin finite element method. The important point is to take $\mathfrak{F}_1$ a domain of interior skewness: with $\mathcal{D} = \partial_x$, we can get away with mere continuity at the nodes, i.e., $\mathfrak{F}_1 = \mathfrak{F}^{\rm pp}(n,0)$. With linear elements ($n = 1$), S and A are tridiagonal and we get a $O(h)$ approximation of $\mathcal{D}v$ on arbitrary grids, where $h = \max(x_{i+1} - x_i)$; $O(h^2)$ approximation on smooth grids (i.e., $x_i = g(ih)$ with g smooth with smooth inverse); and $O(h^4)$ approximation at the nodes on periodic constant-spaced grids (the superconvergence phenomenon [40]). Local, explicit, skew-adjoint finite differences can also be constructed using quite different techniques [7, 17, 23].

3. Example

3.1. **One-way wave equation $\dot u = u_x$.** Consider (semi-)discretizations of

$$\dot u \;=\; \partial_x u \tag{18}$$

on the interval, obtained by globally interpolating point data, computing the derivative of the interpolant, and evaluating this derivative at the grid points.

Solutions of Eq. (18) conserve the energy $\mathcal{H} := \frac{1}{2}\int u^2\,dx$:

$$\dot{\mathcal{H}} \;=\; \int_{-1}^{1} u\dot u\,dx \;=\; \int_{-1}^{1} uu_x\,dx \;=\; \frac{1}{2}u^2(1) - \frac{1}{2}u^2(-1).$$

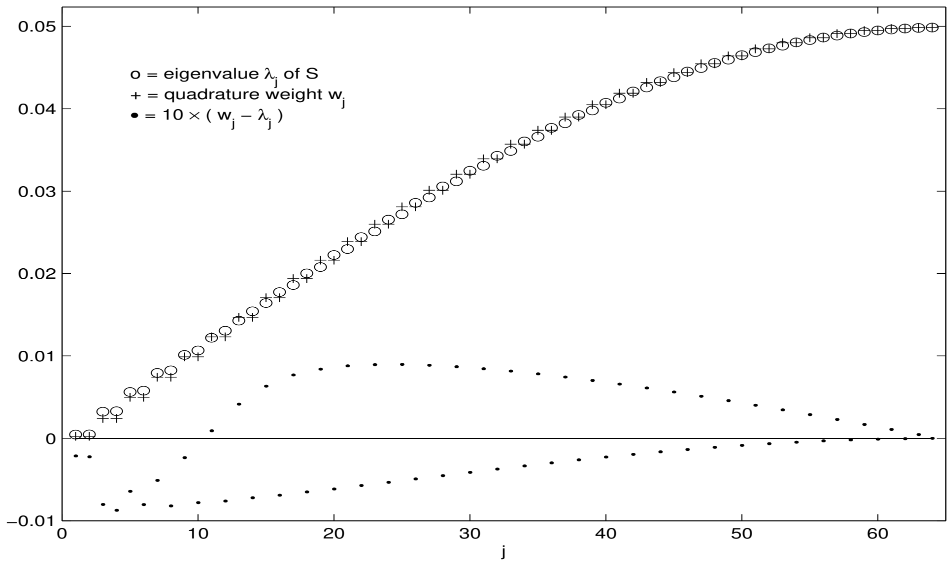

Figure 3: Eigenvalues λ_j of the Chebyshev inner product matrix S compared to the Chebyshev quadrature weights w_j, for $n = 63$. Also shown are the differences $w_j - \lambda_j$ (magnified ten times).

Consequently, with periodic boundary conditions, $\mathcal{H}$ is constant, and otherwise the change in $\mathcal{H}$ is completely determined by the fluxes of u throught the boundary.

Let's first consider periodic boundary conditions, and pseudospectral discretizations based on interpolation with trigonometric basis functions

So the discretization is

$$\dot{\mathbf{u}} \;=\; S^{-1}AS^{-1}\nabla H \;=\; S^{-1}AS^{-1}S\mathbf{u} \;=\; S^{-1}A\mathbf{u} \;=\; D\mathbf{u}$$

just as one would expect.

With periodic boundary conditions, A is antisymmetric, so this is a Hamiltonian discretization, for all basis functions and grids.

For other boundary conditions, the energy H is not conserved, nor does it always evolve appropriately. Consider, for example, the well-posed boundary condition $u(1,t) = f(t)$. With this boundary condition,

$$\dot{\mathcal{H}} = u(1) - u(-1).$$

Take a cardinal basis on a grid which includes the points $x_0 = -1$, $x_n = 1$, and set $u_n(t) = f(t)$, $\dot{\mathbf{u}}_j = (S^{-1}A\mathbf{u})_j$ for $0 \le j < n$, and $H = \frac{1}{2}\mathbf{u}^{\mathrm{T}}S\mathbf{u}$. We set $\boldsymbol{\varepsilon} = \dot{\mathbf{u}} - S^{-1}A\mathbf{u}$ (only ε_n is nonzero). Then

$$\dot{H} \;=\; \mathbf{u}^{\mathrm{T}}S\dot{\mathbf{u}} \;=\; \mathbf{u}^{\mathrm{T}}\,\mathrm{sym}(A)\mathbf{u} + \mathbf{u}^{\mathrm{T}}S\boldsymbol{\varepsilon} \;=\; \frac{1}{2}u_n^2 - \frac{1}{2}u_0^2 + \mathbf{u}^{\mathrm{T}}S\boldsymbol{\varepsilon}.$$

The last term, although usually spectrally small, is not zero: all elements of $\mathbf{u}$ enter the evolution of H. However, if $f(t) \equiv 0$, then u_n and ε vanish and

$$\dot{H} = -\frac{1}{2}u_0^2,$$

exactly as for the PDE. This shows that the discretization is stable in this case.

4. General skew-adjoint differential operators

The above theory is easily applied to more general operators than ∂_x. The required ingredients are a formally skew-adjoint differential operator $\mathcal{D}$ and a finite-dimensional domain of interior skewness $\mathfrak{F}_1$.

4.1. Higher order derivatives. Let $\mathcal{D} = \partial_{xxx}$. We have

$$\int v\mathcal{D}w\,dx = (vw_{xx} - v_x w_x + v_{xx}w) - \int w\mathcal{D}v\,dx \tag{19}$$

so $\mathfrak{F}_1 = C^2([-1,1])$ is a domain of interior skewness. Skew-adjoint finite element methods, for example, can be constructed using C^2 local (Hermite) quintics or C^2 cubic splines. It is known that using less smoothness not only destroys skew-adjointness but also results in suboptimal convergence of finite element methods [40].

The only way to avoid so much smoothness is to use a constant periodic grid. On such a grid, the reflection symmetry of the interpolation process and of the underlying grid (about all grid points as well as the midpoints), together with reflection anti-invariance of the differential operator, guarantee antisymmetry of the differentiation matrix. For example, local, centered, C^0 cubic elements on a constant periodic grid induce a skew-adjoint (actually, antisymmetric) differentiation matrix (although they still have suboptimal convergence); so do standard centered finite differences.

With $\mathfrak{F}_1 = \mathfrak{P}_n([-1,1])$, the Chebyshev pseudospectral differentiation matrix, $D^{(3)}$, is exact on $\mathfrak{F}_1$ and hence does equal the Galerkin matrix in a cardinal basis. It factors as $D^{(3)} = S^{-1}A$. However, the boundary term in (19) is not supported by only two cardinal basis functions as it is for $\mathcal{D} = \partial_x$: the symmetric part of A is dense. Still, $\langle D^{(3)}\mathbf{v}, \mathbf{w}\rangle_S + \langle \mathbf{v}, D^{(3)}\mathbf{w}\rangle_S$ is exact for v, $w \in \mathfrak{F}_1$.

4.2. Nonhomogeneous first order operator. Let $\mathcal{D} = a\partial_x + \partial_x a$, where $a \in C^1([-1,1])$ is a given function. We have

$$\int v\mathcal{D}w\,dx = 2avw - \int w\mathcal{D}v\,dx$$

so that, e.g., $\mathfrak{F}^{\mathrm{pp}}(n,r)$ is a domain of interior skewness, leading to matrix elements of the form $A_{ij} = \int_{-1}^{1} f_i \mathcal{D} f_j\,dx$. If these integrals cannot be evaluated analytically, one can approximate a (by, say, a polynomial) beforehand, or use quadratures to evaluate the integrals to within roundoff error.

It can be seen that implementations can be quite expensive, and that the resulting operator may not be pseudospectral, as it was for $\mathcal{D} = \partial_x$. Nonetheless, the (skew-adjoint) Galerkin approximation can be computed quickly using pseudospectral techniques if fast transforms between cardinal and orthonormal bases are availaible. The key idea is to evaluate $\mathcal{D}v$ exactly for $v \in \mathfrak{F}_1$ and then project orthogonally to $\mathfrak{F}_1$.

Example 5 (**Fast Fourier Galerkin method**)**.** Let $\mathcal{D}(a)$ be linear in a, for example, $\mathcal{D}(a) = a\partial_x + \partial_x a$. Let $\mathfrak{F}_1 = \mathfrak{F}^{\mathrm{trig}}$ be the trigonometric polynomials of degree n, and let a, $v \in \mathfrak{F}_1$. Then $\mathcal{D}(a)v$ is a trigonometric polynomial of degree $2n$, the first n modes of which can be evaluated exactly using antialiasing and Fourier pseudospectral differentiation. Then the approximation whose error is orthogonal to $\mathfrak{F}_1$ is just these first n modes, because $S = I$ in the spectral basis. That is, for skew-adjoint differential operators with linear inhomogeneities, the antialiased pseudospectral method is identical to the Galerkin method, and hence skew-adjoint.

$\mathcal{D}$'s which are polynomial in a can be handled analogously.

Example 6 (**Fast Chebyshev Galerkin method**)**.** Let $\mathfrak{F}_1 = \mathfrak{P}_n$ be the polynomials of degree (no more than) n. Again, let $\mathcal{D}(a)$ be linear in a. As in the previous example, for a and v in $\mathfrak{F}_1$, we evaluate $\mathcal{D}(a)v$ exactly, then orthogonally project onto $\mathfrak{F}_1$. With respect to the cardinal basis corresponding to a Chebyshev grid with $n+1$ points, this can be done in time $O(n \log n)$ as follows:

1. Using an FFT, express a and v as Chebyshev polynomial series of degree n ($O(n \log n)$);
2. Pad with zeros to get Chebyshev polynomial series of formal degree $2n$ ($O(n)$);
3. Transform back to a Chebyshev grid with $2n+1$ points ($O(n \log n)$);
4. Compute the pseudospectral approximation of $\mathcal{D}(a)v$ on the denser grid. $\mathcal{D}(a)v$ being a polynomial of degree $\le 2n$, the corresponding Chebyshev polynomial series is exact ($O(n)$);
5. Convert $\mathcal{D}(a)v$ to a Legendre polynomial series using a fast transform [1] ($O(n \log n)$);
6. Take the first $n+1$ terms ($O(n)$);
7. Evaluate at the points of the original Chebyshev grid (again, [1]) ($O(n \log n)$).

Step 6 produces the Galerkin approximation because the Legendre polynomial basis is orthogonal.

While this algorithm is convoluted, its cost is a small multiple of the cost of standard pseudospectral methods. (Furthermore, we imagine that a fast transforms expert could collapse some of the steps). The algorithm can be adapted to higher degree a's or other $\mathcal{D}$'s by choosing the intermediate Chebyshev grid suitably dense. We do not know how to keep the same asymptotic speed if a grid other than Chebyshev is used.

4.3. **Nonconstant operators $\mathcal{D}(u)$.** We extend the above to the discretization of linear differential operators which depend on the field variable u.

Definition 8. *A nonconstant differential operator*

$$\mathcal{D} : \mathfrak{F}_0 \times \mathfrak{F}_1 \to \mathfrak{F}, \quad \mathcal{D}(u)v \mapsto w,$$

is formally skew-adjoint *with respect to the linear function spaces $\mathfrak{F}_0$ and $\mathfrak{F}_1$ and the inner product space $\mathfrak{F}$ if there exists a functional $b(u, v, w)$, depending on the values of u, v and w at the boundary as well as the boundary values of finitely many of their derivatives, such that*

$$\langle v, \mathcal{D}(u)w \rangle = -\langle w, \mathcal{D}(u)v \rangle + b(u, v, w),\ \forall u \in \mathfrak{F}_0,\ \forall v, w \in \mathfrak{F}_1.$$

Then, $\mathfrak{F}_1$ is called a domain of interior skewness*. If, furthermore, $b(u, v, w) \equiv 0$ for all $u \in \mathfrak{F}_0$ and $v, w \in \mathfrak{F}_1$, then $\mathfrak{F}_1$ is called a* domain of skewness of $\mathcal{D}$*, and $\mathcal{D}$ is said to be* skew-adjoint*.*

We emphasize that the linear operator $\mathcal{D}(u)$ can depend *nonlinearly* on u, e.g., $\mathcal{D}(u)v = u\partial_x(uv)$.

If $\mathcal{D}$ is skew-adjoint for the finite dimensional spaces $\mathfrak{F}_0$, $\mathfrak{F}_1$, then the Galerkin approximation is skew-adjoint for all $u \in \mathfrak{F}_0$.

If $\mathcal{D}$ is linear in u, then $A_{ij}(u) = \langle f_i, \mathcal{D}(u) f_j\rangle$ can be computed with respect to a basis $\{e_k\}$ of $\mathfrak{F}_0$ using the 3-tensor

$$A_{ijk} = \langle f_i, \mathcal{D}(e_k) f_j\rangle\,.$$

For local operators discretized using finite elements, A_{ijk} is local. For Fourier or Chebyshev bases, a fast pseudospectral evaluation can sometimes be found. If $\mathcal{D}$ is fully nonlinear in u, one can evaluate the nonlinear functions in real space, and then interpolate them in $\mathfrak{F}_0$.

Example 7 (KdV operator). The KdV operator

$$\mathcal{D}(u) = 2(u\partial_x + \partial_x u) + \partial_{xxx},$$

which satisfies

$$\langle v, \mathcal{D}(u)w\rangle + \langle w, \mathcal{D}(u)v\rangle = [w_{xx}v - w_x v_x + w v_{xx} + 2uvw],$$

appears in the KdV equation

$$\dot{u} \;=\; \mathcal{D}(u)\frac{\delta\mathcal{H}}{\delta u} \;=\; \mathcal{D}(u)u \;=\; -(2u\partial_x + 2\partial_x u + \partial_{xxx})u \;=\; -(6uu_x + u_{xxx}).$$

Suitable domains of interior skewness are $\mathfrak{F}_0 = \mathfrak{F}^{\mathrm{pp}}(1,0)$, $\mathfrak{F}_1 = \mathfrak{F}^{\mathrm{pp}}(3,2)$. When using local finite element-type basis functions, we can get away with less smoothness in u than in v and w; this can be used to reduce the number of non-zero entries in A without losing the skew-adjoint structure.

5. $S^{-1}A$ factorizations and the Kreiss Matrix Theorem

NICOLAS: PUT YOUR NEW WORK HERE.

6. Discretizing energy conserving PDEs

6.1. Introduction.

Many conservative systems are Hamiltonian; for such systems a Hamiltonian spatial discretization would be attractive. For noncanonical systems, such as the Euler equations, this can be impossible [2]. Nonetheless, if the Hamiltonian structure (i.e., the Poisson bracket) is independent of the field variables, then for the ODEs arising from spatial discretization to be Hamiltonian structure it is sufficient for the discretization of $\mathcal{D}$ to be skew-adjoint.

We consider PDEs of the form

$$\dot{u} = \mathcal{D}(u)\frac{\delta\mathcal{H}}{\delta u}. \tag{20}$$

Such PDEs conserve the energy $\mathcal{H}$. Although all systems of ODEs with an integral can be written in skew-gradient form

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

in which L is antisymmetric [21, 22], it is not clear whether the same is true of all PDEs with an integral. Hamiltonian PDEs are naturally in the form (20) and provide examples [25].

Let $\mathfrak{F}$ be an inner product space. In all our examples, $\mathfrak{F}$ is the space of square integrable functions on the relevant spatial domain, and $\langle u, v\rangle$ is the L^2 inner product $\int uv\,dx$. We use subspaces of $\mathfrak{F}$ which are infinite dimensional when posing the PDE, and finite dimensional when constructing discrete approximations.

The PDE (20) is assumed posed in a subspace $\mathfrak{F}_0$ of $\mathfrak{F}$ which is *natural* for $\mathcal{H}$ in the following sense.

Definition 9 (Naturality). *The subspace* $\mathfrak{F}_0$ *is* natural for $\mathcal{H}$ *if for each* $u \in \mathfrak{F}_0$ *there exists a* $\frac{\delta\mathcal{H}}{\delta u} \in \mathfrak{F}$ *(the variational derivative of* $\mathcal{H}$ *at* u*) such that*

$$\lim_{\varepsilon\to 0} \frac{\mathcal{H}(u+\varepsilon v) - \mathcal{H}(u)}{\varepsilon} = \left\langle v, \frac{\delta\mathcal{H}}{\delta u}\right\rangle,\ \forall v \in \mathfrak{F}. \tag{21}$$

The naturality of $\mathfrak{F}_0$ often follows from the vanishing of the boundary terms, if any, which appear of the first variation of $\mathcal{H}$, together with mild smoothness assumptions (see, e.g., the classic [12]).

Let $\mathcal{H} : \mathfrak{F} \to \mathbb{R}$ be the conserved energy, let $\mathfrak{F}_0$ be natural for $\mathcal{H}$, and let the subspace $\mathfrak{F}_1 \subseteq \mathfrak{F}$ be a domain of skewness for the linear differential operator $\mathcal{D} : \mathfrak{F}_1 \to \mathfrak{F}$.

If $u(t) \in \mathfrak{F}_0$ for all t, we also have $\dot{u}(t) \in \mathfrak{F}_0$. (Nonhomogeneous boundary conditions are not considered here. Generalizations to the nonhomogeneous case, involving affine function spaces, are being developed by the authors.) Accordingly, when $\mathfrak{F}_0$ and $\mathfrak{F}_1$ are infinite dimensional spaces used in describing the PDE, we require that[2]

$$\begin{aligned} \frac{\delta\mathcal{H}}{\delta u}(\mathfrak{F}_0) &\subseteq \mathfrak{F}_1, \text{ and} \\ \mathcal{D}(\mathfrak{F}_1) &\subseteq \mathfrak{F}_0. \end{aligned}$$

The dual composition method uses appropriate finite-dimensional spaces $\mathfrak{F}_0$ and $\mathfrak{F}_1$ to generate spectral, pseudospectral, and finite element discretizations which have discrete energy $H := \mathcal{H}|_{\mathfrak{F}_0}$ as a conserved quantity. The discretization of the differential operator $\mathcal{D}$ is a linear operator $\overline{\mathcal{D}} : \mathfrak{F}_1 \to \mathfrak{F}_0$, and the discretization of the variational derivative $\frac{\delta\mathcal{H}}{\delta u}$ is $\overline{\frac{\delta\mathcal{H}}{\delta u}} \in \mathfrak{F}_1$. Thus

$$\overline{\mathcal{D}}\,\overline{\frac{\delta\mathcal{H}}{\delta u}} : \mathfrak{F}_0 \to \mathfrak{F}_0,$$

as required, for $u \in \mathfrak{F}_0$. In bases of $\mathfrak{F}_0$ and $\mathfrak{F}_1$, the maps have the form

$$\begin{aligned} \overline{\frac{\delta\mathcal{H}}{\delta u}} &: \mathbf{u} \mapsto S^{-1}\nabla H, \\ \overline{\mathcal{D}} &: \mathbf{v} \mapsto S^{-\mathrm{T}}A\mathbf{v}, \\ \overline{\mathcal{D}}\,\overline{\frac{\delta\mathcal{H}}{\delta u}} &: \mathbf{u} \mapsto S^{-\mathrm{T}}AS^{-1}\,\nabla H, \end{aligned}$$

where

(i) ∇H the gradient of $H = \mathcal{H}|_{\mathfrak{F}_0}$ with respect to the coefficients of u in the basis of $\mathfrak{F}_0$;

[2]Exceptions can be sometimes be made due to the existence of implied boundary conditions. In the wave equation $\dot{u} = v$, $\dot{v} = u_{xx}$, $u(\pm 1) = 0$, for example, we also have the implied boundary conditions $v(\pm 1) = u_{xx}(\pm 1) = v_{xx}(\pm 1) = \ldots = 0$. In such cases the operators are in fact applied to subspaces of $\mathfrak{F}_0$ and $\mathfrak{F}_1$, and one then requires, e.g., that $\mathcal{D}(v) \in \mathfrak{F}_0$ for v an element of the appropriate subspace.

(ii) A is the matrix representation of the bilinear form

$$\begin{aligned}\mathcal{A} : \mathfrak{F}_1 \times \mathfrak{F}_1 \to \mathbb{R}\\ \mathcal{A}(v, w) := \langle v, \mathcal{D}w \rangle\,,\end{aligned}$$

induced by the differential operator $\mathcal{D}$ and the inner product, and

(iii) S is the matrix representation of the bilinear form

$$\begin{aligned}\mathcal{S} : \mathfrak{F}_0 \times \mathfrak{F}_1 \to \mathbb{R}\\ \mathcal{S}(u, v) := \langle u, v \rangle\,,\end{aligned}$$

induced by the inner product.

Each of $\overline{\mathcal{D}}$ and $\overline{\frac{\delta\mathcal{H}}{\delta u}}$ is a weighted residual approximation, but in general each uses a space of weight functions different from its space of trial functions. This relies on the following lemma.

Proposition 9. *Let $\mathfrak{F}$ be an inner product space, and $\mathfrak{F}_0$ and $\mathfrak{F}_1$ be subspaces with equal, finite, dimension n. Let $\mathfrak{F}_0$ have basis $\{f_1, \ldots, f_n\}$ and let $\mathfrak{F}_1$ have basis $\{g_1, \ldots, g_n\}$. Then, the following are equivalent.*

(i) $\mathfrak{F}_0^\perp \cap \mathfrak{F}_1 = \{0\}$.

(ii) $\mathfrak{F}_0 \cap \mathfrak{F}_1^\perp = \{0\}$.

(iii) *For every $w \in \mathfrak{F}$, there exists a unique $v \in \mathfrak{F}_1$ such that*

$$\langle u, v \rangle = \langle u, w \rangle\,,\ \forall u \in \mathfrak{F}_0. \tag{22}$$

(iv) *For every $w \in \mathfrak{F}$, there exists a unique $u \in \mathfrak{F}_0$ such that*

$$\langle u, v \rangle = \langle w, v \rangle\,,\ \forall v \in \mathfrak{F}_1.$$

(v) *The $n \times n$ matrix with entries $S_{ij} = \langle f_i, g_j \rangle$ is nonsingular.*

Proof. By symmetry, it is sufficient to show that (i), (iii), and (v) are equivalent.

(i)⇒(v) Suppose that (v) does not hold, that is, S is singular., Then, there exists a non-zero vector $\mathbf{v} = (v_j)$ in the null space of S.

We claim that $v := v_j g_j$ is orthogonal to $\mathfrak{F}_0$, thus contradicting (i). Indeed, let $u = u_i f_i$ be an arbitrary element of $\mathfrak{F}_0$, and let $\mathbf{b} = S\mathbf{v}$, so that $\mathbf{b}$ is null. Writing b_i for the entries of $\mathbf{b}$, we have

$$\langle u, v \rangle \;=\; \langle u_i f_i, v_j g_j \rangle \;=\; u_i \langle f_i, g_j \rangle\, v_j \;=\; u_i b_i \;=\; 0.$$

Consequently, $v \in \mathfrak{F}_0^\perp$.

(v)⇒(iii) First we show that if (v) holds, there exists a v for which (22) holds. Let $b_i = \langle f_i, w \rangle$, $\mathbf{b} = (b_i)$, and $\mathbf{v} = (v_i) = S^{-1}\mathbf{b}$. We claim that $v = v_i g_i$ will do. Indeed, for any $u = u_i f_i$ in $\mathfrak{F}_0$, we have

$$\langle u, v \rangle \;=\; u_i b_i \;=\; u_i \langle f_i, w \rangle \;=\; \langle u_i f_i, w \rangle \;=\; \langle u, w \rangle\,.$$

Now we show that v is unique. If it is not unique, then for some $w \in \mathfrak{F}$ there are two distinct v's satisfying (22). Then, their difference—call it v, for it is in $\mathfrak{F}_1$—is non-zero and satisfies

$$\langle u, v \rangle \;=\; \langle u, w \rangle - \langle u, w \rangle \;=\; 0$$

for all $u \in \mathfrak{F}_0$. Writing $v = v_i g_i$, the vector $\mathbf{v} = (v_i)$ is now seen to be in the null space of S, which contradicts (v).

(iii)⇒(i) A nonzero $v \in \mathfrak{F}_1$ orthogonal to $\mathfrak{F}_0$ has $\langle \mathfrak{F}_0, v \rangle = 0$. But we also have $\langle \mathfrak{F}_0, 0 \rangle = 0$, contradicting (iii). □

Proposition 9 makes it easy to define the discretizations in several equivalent ways. Given $u \in \mathfrak{F}_0$, we want an approximation $\overline{\frac{\delta \mathcal{H}}{\delta u}} \in \mathfrak{F}_1$ of $\frac{\delta \mathcal{H}}{\delta u}$. Given $v \in \mathfrak{F}_1$, we want an approximation $\overline{\mathcal{D}}v \in \mathfrak{F}_0$ of $\mathcal{D}v$. Then we have the following chain of maps:

$$u \in \mathfrak{F}_0 \xrightarrow{\overline{\frac{\delta \mathcal{H}}{\delta u}}} \mathfrak{F}_1 \xrightarrow{\overline{\mathcal{D}}} \mathfrak{F}_0, \tag{23}$$

where $\overline{\mathcal{D}}$ is linear. The approximations must be compatible in order for their composition to have the required skew-gradient structure.

6.2. Discretizing the variational derivative $\delta\mathcal{H}/\delta u$.

Proposition 10. *Let $\mathfrak{F}_0$ be natural for $\mathcal{H}$ and let S be nonsingular. Then for $u \in \mathfrak{F}_0$, the following characterizations define the same unique element $\overline{\frac{\delta \mathcal{H}}{\delta u}} \in \mathfrak{F}_1$, the* discretization of $\frac{\delta \mathcal{H}}{\delta u}$ in $\mathfrak{F}_1$.

(i) *Let $u = u_i f_i$ and let $\mathbf{u} = (u_i)$. Let the discrete energy be $H = \mathcal{H}|_{\mathfrak{F}_0}$, which in the basis has the form*

$$H(\mathbf{u}) = \mathcal{H}(u_i f_i).$$

Let the vector

$$\mathbf{e} \;:=\; (e_i) \;=\; S^{-1}\nabla H.$$

Then,

$$\overline{\frac{\delta \mathcal{H}}{\delta u}} := e_i g_i.$$

(ii) *Let*

$$\mathbf{b} := (b_i) \text{ with } b_i = \left\langle f_i, \frac{\delta \mathcal{H}}{\delta u} \right\rangle,$$

and

$$\mathbf{e} \;:=\; (e_i) \;=\; S^{-1}\mathbf{b}.$$

Then,

$$\overline{\frac{\delta \mathcal{H}}{\delta u}} := e_i g_i.$$

(iii) $\overline{\frac{\delta \mathcal{H}}{\delta u}}$ *is the unique element of $\mathfrak{F}_1$ such that*

$$\left\langle w, \overline{\frac{\delta \mathcal{H}}{\delta u}} \right\rangle = \left\langle w, \frac{\delta \mathcal{H}}{\delta u} \right\rangle, \ \forall w \in \mathfrak{F}_0.$$

(iv) $\overline{\frac{\delta \mathcal{H}}{\delta u}}$ *is the unique element of $\mathfrak{F}_1$ such that*

$$\overline{\frac{\delta \mathcal{H}}{\delta u}} - \frac{\delta \mathcal{H}}{\delta u} \in \mathfrak{F}_0^{\perp}.$$

Proof. Proposition 9 shows that (iii) and (iv) are equivalent. Equivalence of (ii) and (iii) follows from

$$\begin{aligned} \left\langle w, \overline{\frac{\delta \mathcal{H}}{\delta u}} \right\rangle = \left\langle w, \frac{\delta \mathcal{H}}{\delta u} \right\rangle, \ \forall w \in \mathfrak{F}_0 \quad &\Longleftrightarrow \quad w_i S_{ij} e_j \;=\; w_i \left\langle f_i, \frac{\delta \mathcal{H}}{\delta u} \right\rangle \;=\; w_i b_i, \ \forall i \\ &\Longleftrightarrow \quad \mathbf{e} = S^{-1}\mathbf{b}. \end{aligned}$$

Finally, we have to show that $\mathbf{b} = \nabla H$, i.e., that

$$\left\langle f_i, \frac{\delta \mathcal{H}}{\delta u} \right\rangle = \frac{\partial H}{\partial u_i}, \forall i.$$

This follows from the definition of the variational derivative, Eq. (21). Since $\left\langle v, \frac{\delta \mathcal{H}}{\delta u} \right\rangle = \frac{d}{d\varepsilon}\mathcal{H}(u+\varepsilon v)|_{\varepsilon=0}$ for all $v \in \mathfrak{F}$, it is certainly true for $v = f_i \in \mathfrak{F}_0 \subset \mathfrak{F}$. Since $u \in \mathfrak{F}_0$,

$$\frac{d}{d\varepsilon}\mathcal{H}(u+\varepsilon f_i)|_{\varepsilon=0} = \frac{d}{d\varepsilon} H(u_j f_j + \varepsilon f_i) = \frac{\partial H}{\partial u_i}.$$

$\square$

6.3. **Discretizing the linear differential operator $\mathcal{D}$.**

Proposition 11. *Let S be nonsingular. For $v \in \mathfrak{F}_1$, the following characterizations define the same unique element $\overline{\mathcal{D}}v \in \mathfrak{F}_0$, called the discretization of $\mathcal{D}v$. The map $v \mapsto \overline{\mathcal{D}}v$ is linear, with matrix representation $D := S^{-\mathrm{T}}A$.*

(i) *Let $v = v_j g_j$, let $\mathbf{v} = (v_i)$, and let A be the matrix with entries*

$$A_{ij} := \langle g_i, \mathcal{D} g_j \rangle .$$

Let

$$\mathbf{d} := (d_i) = S^{-\mathrm{T}} A \mathbf{v}.$$

Then, $\overline{\mathcal{D}}v := d_i f_i$.

(ii) *Let $\mathbf{b} := (b_i)$ where*

$$b_i = \langle g_i, \mathcal{D}v \rangle ,$$

and let

$$\mathbf{d} = S^{-\mathrm{T}}\mathbf{b} = (d_i).$$

Then,

$$\overline{\mathcal{D}}v := d_i f_i.$$

(iii) *$\overline{\mathcal{D}}v$ is the unique element of $\mathfrak{F}_0$ such that*

$$\langle \overline{\mathcal{D}}v, w \rangle = \langle \mathcal{D}v, w \rangle , \forall w \in \mathfrak{F}_1.$$

(iv) *$\overline{\mathcal{D}}v$ is the unique element of $\mathfrak{F}_0$ such that*

$$\overline{\mathcal{D}}v - \mathcal{D}v \in \mathfrak{F}_1^{\perp}.$$

Proof. As before, Proposition 9 shows that (iii) and (iv) are equivalent. Equivalence of (i), (ii) and (iii) follows from

$$\begin{aligned} \langle \overline{\mathcal{D}}v, w \rangle = \langle \mathcal{D}v, w \rangle , \forall w \in \mathfrak{F}_1 &\iff d_i S_{ij} w_j = v_i \langle \mathcal{D} g_i, g_j \rangle w_j = v_i A^{\mathrm{T}}_{ij} w_j, \forall j \\ &\iff S^{\mathrm{T}}\mathbf{d} = A\mathbf{v} = \mathbf{b}, \quad b_i = \langle g_i, \mathcal{D}v \rangle \\ &\iff \mathbf{d} = S^{-\mathrm{T}} A \mathbf{v}. \end{aligned}$$

Linearity and $D = S^{-\mathrm{T}}A$ follow from (i). $\square$

Parts (iii) of Propositions 10 and 11 show that each discretization is separately of weighted residual type.

Definition 10. $\overline{\mathcal{D}}\overline{\frac{\delta \mathcal{H}}{\delta u}}$ *is called* the dual composition discretization *of* $\mathcal{D}\frac{\delta \mathcal{H}}{\delta u}$.

The matrix representation of $\overline{\mathcal{D}}\overline{\frac{\delta \mathcal{H}}{\delta u}}$ is $S^{-\mathrm{T}}AS^{-1}\nabla H$. The name "dual composition" comes from the dual roles played by $\mathfrak{F}_0$ and $\mathfrak{F}_1$ in defining $\overline{\frac{\delta \mathcal{H}}{\delta u}}$ (Proposition 10) and $\overline{\mathcal{D}}$ (Proposition 11) so that their composition have a skew-gradient structure.

Proposition 12. *If $\mathfrak{F}_1$ is a domain of skewness, the matrix $S^{-\mathrm{T}}AS^{-1}$ is antisymmetric, and the system of ODEs*

$$\dot{\mathbf{u}} = S^{-\mathrm{T}}AS^{-1}\nabla H$$

has H as an integral.

It has been known for a while [14, 17] that Galerkin semidiscretizations of $\dot{u} = \mathcal{D}u$ conserve the restriction of $\langle u, u\rangle$ to $\mathfrak{F}_1$ when $\mathfrak{F}_1$ is a domain of skewness of $\mathcal{A}$. However, a composition of weighted residual discretizations is not necessarily of weighted residual type, nor is a composition of Galerkin discretizations always of Galerkin type (in general, mixed finite element methods are not merely factored "one-step" finite element methods and require their own theory). For this reason, Proposition 12 is more than merely an extension of the above observation to energy functionals other than $\langle u, u\rangle$.

6.4. **When is the dual composition discretization a weighted residual discretization?** We compare the dual composition discretization—which, being a composition, has two "steps"—to two other discretizations of $\mathcal{D}\frac{\delta\mathcal{H}}{\delta u}$:

Definition 11. *The* one-step weighted residual discretization *is the standard weighted residual discretization of $\mathcal{D}\frac{\delta\mathcal{H}}{\delta u}$ with trial functions in $\mathfrak{F}_0$ and weights in $\mathfrak{F}_1$.*

The one-step Galerkin discretization *is the standard weighted residual discretization of $\mathcal{D}\frac{\delta\mathcal{H}}{\delta u}$ with trial functions and weights both in $\mathfrak{F}_0$.*

Thus, one-step Galerkin and one-step weighted residual coincide when $\mathfrak{F}_0 = \mathfrak{F}_1$.

By definition [3, 10, 42], $\overline{\mathcal{D}}\overline{\frac{\delta\mathcal{H}}{\delta u}}$ coincides with the one-step weighted residual discretization of $\mathcal{D}\frac{\delta\mathcal{H}}{\delta u}$ with trial functions in $\mathfrak{F}_0$ and weights in $\mathfrak{F}_1$ if and only if

$$\overline{\mathcal{D}}\overline{\frac{\delta\mathcal{H}}{\delta u}} - \mathcal{D}\frac{\delta\mathcal{H}}{\delta u} \perp \mathfrak{F}_1,$$

and $\overline{\mathcal{D}}\overline{\frac{\delta\mathcal{H}}{\delta u}}$ coincides with the one-step Galerkin discretization of $\mathcal{D}\frac{\delta\mathcal{H}}{\delta u}$ with trial functions in $\mathfrak{F}_0$ and weights in $\mathfrak{F}_0$ if and only if

$$\overline{\mathcal{D}}\overline{\frac{\delta\mathcal{H}}{\delta u}} - \mathcal{D}\frac{\delta\mathcal{H}}{\delta u} \perp \mathfrak{F}_0. \tag{24}$$

Proposition 13. *$\overline{\mathcal{D}}\overline{\frac{\delta\mathcal{H}}{\delta u}}$ coincides with the one-step weighted residual approximation of $\mathcal{D}\frac{\delta\mathcal{H}}{\delta u}$ if and only if*

$$\mathcal{D}\left(\overline{\frac{\delta\mathcal{H}}{\delta u}} - \frac{\delta\mathcal{H}}{\delta u}\right) \perp \mathfrak{F}_1. \tag{25}$$

Proof. We have

$$\overline{\mathcal{D}}\overline{\frac{\delta\mathcal{H}}{\delta u}} - \mathcal{D}\frac{\delta\mathcal{H}}{\delta u} = \mathcal{D}\left(\overline{\frac{\delta\mathcal{H}}{\delta u}} - \frac{\delta\mathcal{H}}{\delta u}\right) + \left(\overline{\mathcal{D}} - \mathcal{D}\right)\overline{\frac{\delta\mathcal{H}}{\delta u}}, \tag{26}$$

the last term being orthogonal to $\mathfrak{F}_1$ by the definition of $\overline{\mathcal{D}}$ (Proposition 11). $\square$

Proposition 14. *$\overline{\mathcal{D}}\overline{\frac{\delta\mathcal{H}}{\delta u}}$ is the standard one-step weighted residual approximation if any of the following conditions holds:*

(i) $\mathcal{D}(\mathfrak{F}_0^\perp) \perp \mathfrak{F}_1$, *or*
(ii) $\frac{\delta\mathcal{H}}{\delta u}$ *is exact, that is,* $\frac{\delta\mathcal{H}}{\delta u} \in \mathfrak{F}_1$.

Proof.

(i) The difference of $\overline{\frac{\delta\mathcal{H}}{\delta u}}$ and $\frac{\delta\mathcal{H}}{\delta u}$ belongs to $\mathfrak{F}_0^\perp$ by definition. Consequently, $\mathcal{D}(\mathfrak{F}_0^\perp) \perp \mathfrak{F}_1$ implies Eq. (25).

(ii) If $\frac{\delta\mathcal{H}}{\delta u} \in \mathfrak{F}_1$, $\overline{\frac{\delta\mathcal{H}}{\delta u}} = \frac{\delta\mathcal{H}}{\delta u}$ by definition, so that the left hand side of (25) vanishes. □

Proposition 15. *$\overline{\mathcal{D}}\overline{\frac{\delta\mathcal{H}}{\delta u}}$ is the standard one-step Galerkin approximation of $\mathcal{D}\frac{\delta\mathcal{H}}{\delta u}$ with trial functions and weights in $\mathfrak{F}_0$ if any of the following conditions hold:*

(i)

$$\mathcal{D}\left(\overline{\frac{\delta\mathcal{H}}{\delta u}} - \frac{\delta\mathcal{H}}{\delta u}\right) \perp \mathfrak{F}_0, \text{ and} \tag{27}$$

$$\left(\overline{\mathcal{D}} - \mathcal{D}\right)\overline{\frac{\delta\mathcal{H}}{\delta u}} \perp \mathfrak{F}_0, \tag{28}$$

or

(ii) *$\overline{\frac{\delta\mathcal{H}}{\delta u}}$ is exact and $\overline{\mathcal{D}}$ is exact on $\mathfrak{F}_1$, that is, $\frac{\delta\mathcal{H}}{\delta u} \in \mathfrak{F}_1$ and $\mathcal{D}(\mathfrak{F}_1) \subseteq \mathfrak{F}_0$.*

Proof.

(i) Use (26) to show that (24) holds.

(ii) The left hand sides of (27) and (28) vanish.

□

Things are simpler if $\mathfrak{F}_0 = \mathfrak{F}_1$.

Proposition 16. *Suppose that $\mathfrak{F}_0 = \mathfrak{F}_1$. Then, $\overline{\mathcal{D}}\overline{\frac{\delta\mathcal{H}}{\delta u}}$ is the one-step Galerkin approximation of $\mathcal{D}\frac{\delta\mathcal{H}}{\delta u}$ if any of the following conditions hold:*

(i) *(27) holds,*

(ii) *$\mathcal{D}(\mathfrak{F}_0^\perp) \perp \mathfrak{F}_1$, or*

(iii) *$\overline{\frac{\delta\mathcal{H}}{\delta u}}$ is exact.*

Proof. If $\mathfrak{F}_0 = \mathfrak{F}_1$, then the one-step weighted residual discretization is actually the one-step Galerkin one.

(i) Proposition 13.

(ii) Proposition 14 part (i).

(iii) Proposition 14 part (ii). □

Note that even if $\mathfrak{F}_0 \neq \mathfrak{F}_1$ and none of $\overline{\mathcal{D}}$ or $\overline{\frac{\delta\mathcal{H}}{\delta u}}$ is Galerkin, it is possible for their composite $\overline{\mathcal{D}}\overline{\frac{\delta\mathcal{H}}{\delta u}}$ to be the one-step Galerkin discretization of $\mathcal{D}\frac{\delta\mathcal{H}}{\delta u}$.

Because $\frac{\delta\mathcal{H}}{\delta u}$ is projected on $\mathfrak{F}_1$ before $\mathcal{D}$ is applied, it is not sufficient that $\mathcal{D}\frac{\delta\mathcal{H}}{\delta u} \in \mathfrak{F}_0$ for $\overline{\mathcal{D}}\overline{\frac{\delta\mathcal{H}}{\delta u}}$ to be the weighted residual or Galerkin approximation. That is, even if the one-step weighted residual discretization is exact, $\overline{\mathcal{D}}\overline{\frac{\delta\mathcal{H}}{\delta u}}$ may not be.

When is it likely that $\frac{\delta\mathcal{H}}{\delta u} \in \mathfrak{F}_1$? An obvious case is when $\mathcal{H} = \frac{1}{2}\langle u, u\rangle$, since the variational derivative is then u itself; $\frac{\delta\mathcal{H}}{\delta u} = u$ belonging to $\mathfrak{F}_1$ for all u's in $\mathfrak{F}_0$ is then equivalent to $\mathfrak{F}_0 = \mathfrak{F}_1$. But in this case, $\mathcal{D}\frac{\delta\mathcal{H}}{\delta u} = \mathcal{D}u$, and the discretization of $\mathcal{D}$ is obviously the Galerkin one! A less obvious situation is when $\mathcal{H} := \int u_x^2$, and in such contexts, the above propositions may be useful.

When do we have $\mathcal{D}(\mathfrak{F}_0^\perp) \perp \mathfrak{F}_1$? This condition is harder to satisfy because it involves the operator $\mathcal{D}$, the inner product $\langle\,,\rangle$, and the large space $\mathfrak{F}_0^\perp \subset \mathfrak{F}$. However, suppose that $\mathfrak{F}_0 = \mathfrak{F}_1$ and $\overline{\mathcal{D}}$ is exact, i.e., $\mathcal{D}(\mathfrak{F}_0) \subset \mathfrak{F}_1$. Then, if $\mathcal{D}$ and orthogonal projection to $\mathfrak{F}_0$ commute, we have the orthogonal splitting

$$\mathcal{D}(\mathfrak{F}_0 \oplus \mathfrak{F}_0^\perp) = \mathcal{D}(\mathfrak{F}_0) \oplus \mathcal{D}(\mathfrak{F}_0^\perp);$$

consequently, $\mathcal{D}(\mathfrak{F}_0^\perp) \perp \mathfrak{F}_1$. This can happen if $\mathfrak{F}$ has an orthogonal basis of eigenfunctions of $\mathcal{D}$, an example being the eigenfunctions e^{ijx} of $\mathcal{D} = \partial_x^n$. So this situation is not uncommon in Fourier spectral methods; we give an example in Section 7.2.

The solution of a system of PDEs obtained using bases of finite elements with the dual composition method is often identical to the solution of a mixed finite element method discretization [4].

6.5. **Skew-adjointness of the discrete derivative.** Let $\mathfrak{F}_1$ be a domain of skewness for $\mathcal{D}$. Since the domain and range of $\overline{\mathcal{D}} : \mathfrak{F}_0 \to \mathfrak{F}_1$ may be different, it does not make sense to ask if it is skew-adjoint. Indeed, its matrix representation $D = S^{-\mathrm{T}}A$ is not in general skew-adjoint when $\mathfrak{F}_0 \neq \mathfrak{F}_1$ (usually, S is not symmetric when $\mathfrak{F}_0 \neq \mathfrak{F}_1$). Nonetheless, the dual composition method always yields a skew-adjoint discretization of $\mathcal{D}$, as we now show.

By construction (see Eq. (23)), the composite $\overline{\mathcal{D}}\overline{\frac{\delta\mathcal{H}}{\delta u}}$ maps $\mathfrak{F}_0$ to itself. If $\mathcal{H}(u) = \frac{1}{2}\langle u, u\rangle$, then $\frac{\delta\mathcal{H}}{\delta u} = u$, and so $\mathcal{D}\frac{\delta\mathcal{H}}{\delta u} = \mathcal{D}u$, and its discretization $\widetilde{D} := \overline{\mathcal{D}}\overline{\frac{\delta\mathcal{H}}{\delta u}}$ may be examined for skew-adjointness. We do this in matrix form.

As always, let $\{f_i\}$ be a basis of $\mathfrak{F}_0$ and $\{g_j\}$ be a basis of $\mathfrak{F}_1$. Recall that S is the matrix with entries $\langle f_i, g_j\rangle$. Let $\widetilde{S}$ be the matrix of the inner product $\langle\, ,\rangle$ with respect to the basis $\{f_i\}$, i.e.,

$$\widetilde{S}_{i,j} := \langle f_i, f_j\rangle\,.$$

By Proposition 2, $\widetilde{S}$ is symmetric positive definite. In general, $\widetilde{S}$ and S are not equal unless the bases $\{f_i\}$ and $\{g_j\}$ are identical.

With $\mathcal{H}(u) = \frac{1}{2}\langle u, u\rangle$, we have that

$$H(\mathbf{u}) \;:=\; \mathcal{H}(u_i f_i) \;=\; \frac{1}{2}\mathbf{u}^{\mathrm{T}}\widetilde{S}\mathbf{u},$$

so that $\nabla H = \widetilde{S}\mathbf{u}$ and

$$\widetilde{D}\mathbf{u} \;=\; S^{-\mathrm{T}}AS^{-1}\nabla H(\mathbf{u}) \;=\; S^{-\mathrm{T}}AS^{-1}\widetilde{S}\mathbf{u}.$$

This motivates the following:

Definition 12. *The dual composition differentiation matrix $\widetilde{D}$ is*

$$\widetilde{D} := S^{-\mathrm{T}}AS^{-1}\widetilde{S}.$$

If $\{f_i\} = \{g_i\}$, then $S = \widetilde{S}$, and $\widetilde{D}$ is exactly the Galerkin differentiation matrix; in general, however, it is not. The factor $S^{-1}\widetilde{S}$ of the dual composition differentiation matrix is a discretization of the identity map: By Proposition 10, for $u = u_i f_i \in \mathfrak{F}_0$, $S^{-1}\widetilde{S}\mathbf{u}$ is the vector representation of the unique v in $\mathfrak{F}_1$ with the property that $v - u \in \mathfrak{F}_0^\perp$, that is, it is the projection of $u \in \mathfrak{F}_0$ onto $\mathfrak{F}_1$.

Proposition 17. *Let*

$$\langle \mathbf{v}, \mathbf{w}\rangle_{\widetilde{S}} := \mathbf{v}^{\mathrm{T}}\widetilde{S}\mathbf{w}$$

(recall that $\widetilde{S}$ is symmetric positive definite) and $\mathfrak{F}_1$ be a domain of skewness of $\mathcal{D}$. Then, $\widetilde{D}$ is skew-adjoint with respect to $\langle\, ,\rangle_{\widetilde{S}}$.

Proof. We are given that A is antisymmetric. Then we have

$$\widetilde{D} \;=\; S^{-\mathrm{T}}AS^{-1}\widetilde{S} \;=\; \widetilde{S}^{-1}(\widetilde{S}S^{-\mathrm{T}}AS^{-1}\widetilde{S}),$$

where the factor in parenthesis is antisymmetric. □

Finally, note that when $\widetilde{D}$ is used to construct skew-gradient systems, they are of the form $\widetilde{D}\widetilde{S}^{-1}\nabla H$, which is equal to the dual composition discretization $DS^{-1}\nabla H$. We do not get new methods from $\widetilde{D}$, but it does illustrate where the skew-adjointness is hiding.

7. Examples

7.1. **One-way wave equation $\dot u = u_x$.** In Hamiltonian form,

$$\dot u \;=\; \partial_x u \;=\; \mathcal{D}\frac{\delta\mathcal{H}}{\delta u},$$

where

$$\mathcal{D} = \partial_x, \quad \mathcal{H} = \frac{1}{2}\int_{-1}^{1} u^2\, dx.$$

We have

$$\frac{d}{d\varepsilon}\mathcal{H}(u+\varepsilon v) = \int_{-1}^{1} uv\, dx,$$

so $\frac{\delta\mathcal{H}}{\delta u} = u$ with no natural boundary conditions: any space of square integrable functions is natural for $\mathcal{H}$. We have $\langle u, \mathcal{D}v\rangle + \langle v, \mathcal{D}u\rangle = [uv]$, so any space of continuous, piecewise C^1 functions is a domain of interior skewness. Consider the domain $[-1,1]$ and let $\mathfrak{F}_0 = \mathfrak{F}_1 = \{f_1,\ldots,f_n\}$, $f_i \in C^0([-1,1])$. The discrete Hamiltonian is

$$H(\mathbf{u}) \;=\; \frac{1}{2}\int_{-1}^{1} (u_i f_i)^2\; dx \;=\; \frac{1}{2}\mathbf{u}^{\mathrm{T}} S\mathbf{u}.$$

So the discretization is

$$\dot{\mathbf{u}} \;=\; S^{-1}AS^{-1}\nabla H \;=\; S^{-1}AS^{-1}S\mathbf{u} \;=\; S^{-1}A\mathbf{u} \;=\; D\mathbf{u}$$

just as one would expect.

With periodic boundary conditions, A is antisymmetric, so this is a Hamiltonian discretization, for all basis functions and grids.

For other boundary conditions, the energy H is not conserved, nor does it always evolve appropriately. Consider, for example, the well-posed boundary condition $u(1,t) = f(t)$. With this boundary condition,

$$\dot{\mathcal{H}} \;=\; \int_{-1}^{1} u\dot u\, dx \;=\; \int_{-1}^{1} uu_x\, dx \;=\; \frac{1}{2}u^2(1) - \frac{1}{2}u^2(-1).$$

Take a cardinal basis on a grid which includes the points $x_0 = -1$, $x_n = 1$, and set $u_n(t) = f(t)$, $\dot{\mathbf{u}}_j = (S^{-1}A\mathbf{u})_j$ for $0 \le j < n$, and $H = \frac{1}{2}\mathbf{u}^{\mathrm{T}}S\mathbf{u}$. We set $\boldsymbol{\varepsilon} = \dot{\mathbf{u}} - S^{-1}A\mathbf{u}$ (only ε_n is nonzero). Then

$$\dot H \;=\; \mathbf{u}^{\mathrm{T}}S\dot{\mathbf{u}} \;=\; \mathbf{u}^{\mathrm{T}}\,\mathrm{sym}(A)\mathbf{u} + \mathbf{u}^{\mathrm{T}}S\boldsymbol{\varepsilon} \;=\; \frac{1}{2}u_n^2 - \frac{1}{2}u_0^2 + \mathbf{u}^{\mathrm{T}}S\boldsymbol{\varepsilon}.$$

The last term, although usually spectrally small, is not zero: all elements of $\mathbf{u}$ enter the evolution of H. However, if $f(t) \equiv 0$, then u_n and ε vanish and

$$\dot{H} \;=\; -\frac{1}{2}u_0^2,$$

exactly as for the PDE. This shows that the discretization is stable in this case.

This standard method is not exactly a dual composition discretization when the boundary conditions are not periodic because we are not projecting the derivative of the interpolating polynomial onto the space of functions satisfying the boundary conditions. Nonetheless, the factorization $D = S^{-1}A$ is very informative. It is not surprising that energy estimates cannot be obtained in general for the boundary condition $u_n = f(t)$, for the solution u is properly an element of an affine, not a linear space. In the example above, we crudely enforced the boundary conditions and destroyed the structure of the method unless $f(t) \equiv 0$. We are currently studying special approximations that respect such affine structures and comparing them with other discretizations in which the boundary conditions are enforced weakly [5, 9].

7.2. **The Korteweg-de Vries equation.** The famous KdV equation

$$\dot{u} + 6uu_x + u_{xxx} = 0$$

has features which can be used to illustrate various properties of conservative discretization. In its first Hamiltonian form,

$$\dot{u} = \mathcal{D}_1\frac{\delta(\mathcal{H}_0+\mathcal{H}_1)}{\delta u}, \quad \mathcal{D}_1 = \partial_x, \quad \mathcal{H}_0 = \int -u^3\,dx, \quad \mathcal{H}_1 = \frac{1}{2}\int u_x^2\,dx,$$

it has a nonlinear variational derivative and a constant differential operator. In its second Hamiltonian form,

$$\dot{u} = \mathcal{D}_2\frac{\delta \mathcal{H}_2}{\delta u}, \quad \mathcal{D}_2 = -(2u\partial_x + 2\partial_x u + \partial_{xxx}), \quad \mathcal{H}_2 = \frac{1}{2}\int u^2\,dx,$$

$\frac{\delta\mathcal{H}_2}{\delta u} = u$ is linear while the differential operator depends (linearly) on the field variable u. Boundary conditions which make domains natural or skew are different for the two Hamiltonian forms; we consider only periodic boundary conditions so that these complications are not present.

$\mathcal{D}_1 = \partial_x$ is discussed in Section 2. Its approximation is exact on Fourier polynomials $\mathfrak{F}^{\rm trig}$, but not on piecewise polynomials $\mathfrak{F}^{\rm pp}$. We have

$$\frac{\delta\mathcal{H}_1}{\delta u} = -3u^2 - u_{xx},$$

so any space of C^1 and piecewise C^2 functions is natural for $\mathcal{H}$. We consider first the approximation of the cubic term in the Hamiltonian. For a direct evaluation,

$$H_0(\mathbf{u}) \;:=\; \mathcal{H}_0(u_i f_i) \;=\; F_{ijk}u_iu_ju_k, \quad F_{ijk} := \int -f_if_jf_k\,dx,$$

The weighted residual approximation is

$$\overline{\frac{\delta\mathcal{H}_0}{\delta u}} = 3S_{ij}^{-1}F_{jkl}u_ku_l.$$

In general, these quadratic forms are not equal to u_i^2.

We now specialize to $\mathfrak{F}_0 \;=\; \mathfrak{F}_1 \;=\; \mathfrak{F}^{\rm trig}$. In this case the Galerkin approximation of $\frac{\delta\mathcal{H}_0}{\delta u}$ can be evaluated in less direct ways: $\overline{\frac{\delta\mathcal{H}_0}{\delta u}}$ is the orthogonal projection to

$\mathfrak{F}_0$ of $\frac{\delta\mathcal{H}_0}{\delta u} = 3u^2$ which can be computed by multiplying out the Fourier series and dropping all but the first n modes. This is exactly what is computed by antialiasing.

For the second part of the first Hamiltonian, we have

$$H_1(\mathbf{u}) \;:=\; \mathcal{H}_1(u_i f_i) \;=\; \frac{1}{2}\mathbf{u}^{\mathrm{T}}B\mathbf{u}, \quad B_{ij} \;=\; \int f_i' f_j'\,dx \;=\; -\int f_i f_j''\,dx,$$

the last equality holding because of the periodic boundary conditions. Then

$$\overline{\frac{\delta\mathcal{H}_1}{\delta u}} = S^{-1}B\mathbf{u}.$$

Let the pseudospectral second derivative be $D^{(2)}$; in a cardinal basis, we have $B = -SD^{(2)}$, so

$$\overline{\frac{\delta\mathcal{H}_1}{\delta u}} = -D^{(2)}\mathbf{u},$$

as expected for $\frac{\delta\mathcal{H}_1}{\delta u} = -u_{xx}$. The weighted residual and pseudospectral approximations are identical in this case.

We now let $v = \overline{\frac{\delta(\mathcal{H}_0+\mathcal{H}_1)}{\delta u}}$ and construct $\overline{\mathcal{D}}_1 v$. This is just $\mathcal{D}_1 v = v_x$, since differentiation is exact in $\mathfrak{F}^{\mathrm{trig}}$. Moreover, since $\mathcal{D}_1$ is constant, the discretization is a Hamiltonian system. Since $\overline{\mathcal{D}}_1$ is exact on constants, the approximation also preserves the restriction of the Casimir $\mathcal{C} = \int u\,dx$ of $\mathcal{D}_1$, to $\mathfrak{F}^{\mathrm{trig}}$.

How does $\overline{\mathcal{D}}_1\overline{\frac{\delta(\mathcal{H}_0+\mathcal{H}_1)}{\delta u}}$ compare with the one-step Galerkin approximation of $\dot{u} = \mathcal{D}_1\frac{\delta(\mathcal{H}_0+\mathcal{H}_1)}{\delta u}$ in $\mathfrak{F}^{\mathrm{trig}}$? They are in fact the same, for differentiation does not map high modes to lower modes, that is, $\mathcal{D}_1(\mathfrak{F}^{\mathrm{trig}\perp}) \perp \mathfrak{F}^{\mathrm{trig}\perp}$ and Proposition 16 part (ii) applies.

In the second Hamiltonian form, $\mathcal{H}_2 = \frac{1}{2}\int u^2\,dx$ and

$$H_2 = \frac{1}{2}\mathbf{u}^{\mathrm{T}}S\mathbf{u}, \quad \overline{\frac{\delta\mathcal{H}_2}{\delta u}} = S^{-1}\nabla H_2 = \mathbf{u}.$$

Thus, the Galerkin approximation of $\mathcal{H}_2$ is exact, so that Proposition 16 part (iii) implies that the composition $\overline{\mathcal{D}}_2\overline{\frac{\delta\mathcal{H}_2}{\delta u}}$ *also* coincides with the one-step Galerkin approximation. The second differential operator, $\mathcal{D}_2$, can be applied to any $v \in \mathfrak{F}^{\mathrm{trig}}$ by computing the derivatives exactly and the quadratic terms exactly by antialiasing. Its approximation $\overline{\mathcal{D}}_2$ is not a Hamiltonian operator, but we still generate a skew-gradient system with integral H_2. Thus in this, admittedly unusual, case, the standard Galerkin approximation coincides with the standard antialiased pseudospectral method and has three conserved quantities, $H_0 + H_1$, H_2, and $\mathcal{C}|_{\mathfrak{F}^{\mathrm{trig}}}$.

In spaces other than $\mathfrak{F}^{\mathrm{trig}}$ the situation is a little different. Consider, for example, $\mathfrak{F}_0 = \mathfrak{F}_1 = \mathfrak{F}^{\mathrm{pp}}(n,r)$, the C^r piecewise polynomials of degree n, as in a finite element method. In the first Hamiltonian form, we need $r \geq 1$ to ensure that $\mathfrak{F}_0$ is natural for $\mathcal{H}_1$; in the second form, naturality is no restriction, but we need $r \geq 2$ to ensure that $\mathfrak{F}_1$ is a domain of skewness. The skew-gradient system arising from discretizing the first Hamiltonian form is still Hamiltonian with integral $H_0 + H_1$ and Casimir $C = u_i \int f_i\,dx$, but because $\overline{\mathcal{D}}_1$ does not commute with projection to $\mathfrak{F}_1$, it is *not* a standard Galerkin method. To implement this method, the tensors F_{ijk} and B_{ij} will generally have to be evaluated and stored.

In the second Hamiltonian form, $\frac{\delta\mathcal{H}_2}{\delta u} = u$ is still exact, so the resulting skew-gradient system does coincide with the standard Galerkin method. However, the discretization, this time, is not Hamiltonian.

Table 1 summarizes the results.

Table 1: Features of dual composition discretizations of the KdV equation on the circle, using the function spaces $\mathfrak{F}^{\rm trig} :=$ trigonometric polynomials or $\mathfrak{F}^{\rm pp}(r) := C^r$ piecewise polynomials (of suitable degree).

$\mathfrak{F}_0 = \mathfrak{F}_1$	**First Hamiltonian form** $\mathcal{D} = \partial_x$ $\mathcal{H} = \int (\frac{1}{2}u_x^2 - u^3)\, dx$	**Second Hamiltonian form** $\mathcal{D} = -2(u\partial_x + \partial_x u) - \partial_{xxx}$ $\mathcal{H} = \int \frac{1}{2}u^2\, dx$
$\mathfrak{F}^{\rm trig}$	skew, natural Hamiltonian Galerkin	skew, natural conservative, not Hamiltonian Galerkin = antialiased
$\mathfrak{F}^{\rm pp}(0)$	skew, not natural not conservative not Galerkin	natural, not skew not conservative Galerkin
$\mathfrak{F}^{\rm pp}(1)$	skew, natural Hamiltonian not Galerkin	natural, not skew not conservative Galerkin
$\mathfrak{F}^{\rm pp}(2)$	skew, natural Hamiltonian not Galerkin	skew, natural conservative, not Hamiltonian Galerkin

7.3. **The inhomogeneous wave equation.** Even when natural boundary conditions are present in $\mathfrak{F}_0$, it is not always reasonable to include them in $\mathfrak{F}_1$ as well. We present a simple example in which it is necessary to take $\mathfrak{F}_0 \neq \mathfrak{F}_1$.

Consider the wave equation

$$\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} a(x)p \\ q_{xx} \end{pmatrix} \tag{29}$$

on the interval $[-1,1]$ with position dependent wave speed $\sqrt{a}$ and boundary conditions

$$q_x(\pm 1, t) = 0. \tag{30}$$

This is a canonical Hamiltonian system with generalized coordinate q and conjugate momentum p: Let

$$\mathcal{D} := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

and

$$\mathcal{H} := \frac{1}{2}\int_{-1}^{1} \left(ap^2 + q_x^2\right)\, dx.$$

Then,

$$\frac{\delta\mathcal{H}}{\delta q} = -q_{xx}, \quad \frac{\delta\mathcal{H}}{\delta p} = ap,$$

so that $\mathcal{D}\frac{\delta\mathcal{H}}{\delta(q,p)}$ is the right hand side of (29). Note that (i) the boundary conditions (30) are natural for $\mathcal{H}$, and (ii) $\mathcal{D}$ is skew-adjoint with respect to the direct sum

L^2 inner product

$$\langle (p,q),(\tilde{p},\tilde{q})\rangle := \int_{-1}^{1} p\tilde{p}\,dx + \int_{-1}^{1} q\tilde{q}\,dx \tag{31}$$

(no boundary conditions are required for $\mathcal{D}$ to be skew-adjoint).

We now consider the choice of the trial function spaces $\mathfrak{F}_0$ and $\mathfrak{F}_1$. The (natural) boundary conditions (30) are included in the solution space $\mathfrak{F}_0$:

$$\mathfrak{F}_0 \;=\; \mathfrak{F}_0^{(q)} \times \mathfrak{F}_0^{(p)} \;:=\; \{q \in \mathfrak{P}_{n+2} : q_x(-1) = q_x(1) = 0\} \times \mathfrak{P}_n.$$

Since the discrete variational derivative $\overline{\frac{\delta\mathcal{H}}{\delta(q,p)}}$ is to belong to $\mathfrak{F}_1$, $\overline{\frac{\delta\mathcal{H}}{\delta q}}$ would not be a good approximation if we were to include (30) in $\mathfrak{F}_1$ as well, for this would be to enforce $(-q_{xx})_x(\pm 1) = 0$ which is not in general satisfied by q.[3] Thus, we do not impose any boundary conditions on $\mathfrak{F}_1$:

$$\mathfrak{F}_1 \;=\; \mathfrak{F}_1^{(q)} \times \mathfrak{F}_1^{(p)} \;:=\; \mathfrak{P}_n \times \mathfrak{P}_n.$$

Both $\mathfrak{F}_0$ and $\mathfrak{F}_1$ have dimension $2(n+1)$. Before proceeding, we need to show that the S matrix is invertible. By Proposition 9 it is sufficient to prove the following.

Proposition 18. $\mathfrak{F}_0 \cap \mathfrak{F}_1^{\perp} = \{0\}$ *(with respect to the direct sum inner product (31)).*

Proof. $\mathfrak{F}_0^{(p)} \cap \mathfrak{F}_1^{(p)\perp} = \{0\}$ since $\mathfrak{F}_0^{(p)} = \mathfrak{F}_1^{(p)}$.

We now show that $\mathfrak{F}_0^{(q)} \cap \mathfrak{F}_1^{(q)\perp} = \{0\}$. If q is a polynomial of degree at most $n+1$ which is orthogonal to all polynomials with degree at most $n-1$, then it is orthogonal to q_{xx}. If, in addition, $q_x(\pm 1) = 0$, then

$$\int_{-1}^{1} q_x q_x\,dx \;=\; [q q_x]_{-1}^{1} - \int_{-1}^{1} q q_{xx}\,dx \;=\; 0.$$

Consequently, q_x is identically zero, and q is constant. Since q must then be orthogonal to itself, $q = 0$. $\square$

A computer implementation of the method requires chosing bases (cardinal or not) for $\mathfrak{F}_0$ and $\mathfrak{F}_1$. Following is a description in terms of the Chebyshev extreme points $\{x_i\}_{i=0}^{n}$, using the standard cardinal basis for $\mathfrak{F}_0^{(p)} = \mathfrak{F}_1^{(q)} = \mathfrak{F}_1^{(p)}$. For $\mathfrak{F}_0^{(q)}$, we can start with the basis

$$\begin{aligned} f_0(x) &= 1 \\ f_j(x) &= \frac{T_j(x)}{j^2} - \frac{T_{j+2}(x)}{(j+2)^2},\; j = 1,\ldots,n, \end{aligned}$$

chosen so that $f_j'(\pm 1) = 0$ evaluate the inner product matrices analytically, and transform to a cardinal basis as in Section 2.5. (Alternatively, one could enforce the boundary conditions on the basis functions by adding to T_j suitable multiples of T_1 and T_2 instead of T_{j+2} [3, 14].)

[3] Sometimes $(-q_{xx})_x(\pm 1) = 0$ *is* satisfied by solutions of the PDE: from (29)-(30), we have the implied boundary conditions

$$0 \;=\; \dot{q}_x(\pm 1) \;=\; (ap)_x(\pm 1).$$

Differentiating again with respect to time:

$$0 \;=\; (a\dot{p})_x(\pm 1) \;=\; (aq_{xx})_x(\pm 1) \;=\; a_x(\pm 1)q_{xx}(\pm 1) + a(\pm 1)q_{xxx}(\pm 1).$$

If $a_x(\pm 1) = 0$ and $a(\pm 1) \neq 0$—if, for example, the medium is homogeneous—then $q_{xxx}(\pm 1) = 0$ and we can take $\mathfrak{F}_1 = \mathfrak{F}_0$.

We now outline the four steps in computing the discretization, indicating why the method has spectral accuracy. Let us suppose we are given the grid values of $(q,p) \in \mathfrak{F}_0$.

(i) $\frac{\delta H}{\delta q} = -q_{xx} \in \mathfrak{F}_1^{(q)}$, so $\overline{\frac{\delta H}{\delta q}} = -q_{xx}$. This can be computed using standard Chebyshev pseudospectral differentiation. For nonlinear equations, the weighted residual approximation to $\frac{\delta \mathcal{H}}{\delta q}$ with test functions in $\mathfrak{F}_1^{(q)}$ and weights in $\mathfrak{F}_0^{(q)}$ is still exact when $\frac{\delta \mathcal{H}}{\delta q} \in \mathfrak{P}_n$, hence it is spectrally accurate in general.

(ii) $\frac{\delta H}{\delta p} = ap \notin \mathfrak{F}_1^{(p)}$. However, the weights and trial functions are the same in this component, so we compute a standard Galerkin approximation $\overline{ap} \in \mathfrak{P}_n$ of ap. This can be done with the fast Chebyshev-Galerkin method of Example 6.

(iii) The first component of $\mathcal{D}\delta\mathcal{H}$ is $v := \overline{ap}$ which is not in $\mathfrak{F}_0^{(q)}$ because it does not have zero derivatives at the boundary. We need to compute its weighted residual approximation with test functions in $\mathfrak{F}_0^{(q)}$ and trial functions in $\mathfrak{F}_1^{(q)} = \mathfrak{P}_n$. In coordinates, this is $S^{-\mathrm{T}}S^{(2)}\mathbf{v}$ where $S^{(2)}$ is the (symmetric) matrix of the inner product $\langle\, ,\rangle|_{\mathfrak{P}_n\times\mathfrak{P}_n}$. In general, the matrix $S^{-\mathrm{T}}S^{(2)}$ is computed and kept in storage. When evaluated on solutions of the PDE, it is a spectrally accurate approximation because ap satisfies the implied boundary condition $(ap)_x(\pm 1) = 0$, so there is no loss of accuracy in approximating it in $\mathfrak{F}_0^{(q)}$.

(iv) The second component of $\mathcal{D}\delta H$ is $w := q_{xx} \in \mathfrak{F}_0^{(p)}$, so it does not need to be approximated.

Unlike standard Galerkin or pseudospectral methods, this method is conservative and Hamiltonian (by design) and enforces the implied boundary condition $(ap)_x(\pm 1) = 0$.

We plan to report on numerical tests exploring the effects of these changes in the future, particularly when the initial data does not satisfy the implied boundary conditions.

8. Quadrature of Hamiltonians

8.1. The problem. The discretization

$$\dot{\mathbf{u}} = S^{-\mathrm{T}}AS^{-1}\nabla H$$

is skew-gradient—and, when $\mathcal{D}$ does not depend on u, Hamiltonian—when appropriate boundary conditions make A antisymmetric. Computing $\nabla H = \nabla\mathcal{H}(u_j f_j)$ in closed form, however, is not always possible. If the elements of $\mathfrak{F}_0$ are polynomial or piecewise polynomial, and $\mathcal{H}$ is polynomial in u and its derivatives, then ∇H can be evaluated (exactly) using quadratures on a secondary grid chosen so that the quadratures are exact. For other functions, even this is not possible: Consider $\mathcal{H} = \int h(u)\,dx$, with $\frac{\delta\mathcal{H}}{\delta u} = h'(u)$; then,

$$\frac{\partial H}{\partial u_k} = \int h'(u_j f_j) f_k\,dx,$$

which usually cannot be computed exactly. (It could be approximated to within roundoff error using quadratures, but this brings in new realms of computational complexity.)

We would like to approximate $\mathcal{H}$ itself by quadratures in real space, analogous to a pseudospectral method. However, *even if the discrete H and its gradient* *are spectrally accurate approximations, they cannot always be used to construct spectrally accurate Hamiltonian discretizations,* as we now illustrate.

For simplicity, we now assume that the basis of $\mathfrak{F}$ is cardinal. Define the weights $w_j = \int f_j\,dx$ and $W = \mathrm{diag}(\mathbf{w})$, so that quadrature in $\mathfrak{F}$ is

$$\int u_j f_j\,dx = u_j w_j.$$

In the case $\mathcal{H} = \int h(u)\,dx$, the quadrature Hamiltonian H_q is defined by

$$H_q \;=\; h(u_j)w_j \;=\; \mathbf{w}^{\mathrm{T}}h(\mathbf{u})$$

with gradient

$$\nabla H_q = Wh'(\mathbf{u}).$$

This is equal to the quadrature of the gradient of the exact Hamiltonian $H = \int h(u_j f_j)$. (Basis independence, retained in $\mathcal{A}$ and $\mathcal{S}$, is now lost, for quadratures depend on the grid.) So far, everything is as accurate as the quadratures. However, $S^{-1}\nabla H_q = S^{-1}Wh'(\mathbf{u})$ *is not an accurate approximation of* $\frac{\delta\mathcal{H}}{\delta u} = h'(\mathbf{u})$.

(Why is this? Consider the cardinal functions on the Chebyshev extreme points $x_i = -\cos(i\pi/n)$. Spectral accuracy in $S^{-1}W\mathbf{v} \approx \mathbf{v}$ requires that the eigenvalues of $S^{-1}W$ be spectrally close to 1. Numerically, it appears that this is true for all but three eigenvalues. One of the "spurious" eigenvectors is a sawtooth and two are supported near the endpoints. These limit the accuracy of $S^{-1}W\mathbf{v}$ as an approximation of $\mathbf{v}$ to $O(n^{-2})$, that is, spectral accuracy has been lost.)

Nor can we use the accurate approximation $h'(\mathbf{u})$, for $Sh'(\mathbf{u})$ is not, in general, a gradient, which it must be if $\dot{\mathbf{u}} = S^{-\mathrm{T}}AS^{-1}(Sh'(\mathbf{u}))$ is to be a skew-gradient system.

We found only one non-periodic case in which quadrature yields an accurate Hamiltonian discretization for any $\mathcal{D}$ and $\mathcal{H}$.

8.2. **Legendre zeros grid.** Let $\mathfrak{F}_0 = \mathfrak{F}_1$ be the polynomials of degree n, with a basis $\{f_j\}$ of cardinal functions at the Legendre points, namely the roots of the Legendre polynomial of degree $n+1$. These are orthogonal, so S is a diagonal matrix, but more is true: $S = W$, the matrix of quadrature weights—the Gaussian weights in this case [3]. This is because quadrature at the Legendre points is exact for polynomials of degree $\le 2n+1$, so

$$S_{ii} \;=\; \int f_i f_i\,dx \;=\; f_i(x_j)^2 w_j \;=\; w_i$$

since f_i^2 is a polynomial of degree precisely $2n$.

Therefore, $\frac{\delta\mathcal{H}}{\delta u} \approx W^{-1}\nabla H_q = S^{-1}\nabla H_q$ and we get skew-gradient, spectrally accurate discretizations of the form

$$\dot{\mathbf{u}} = S^{-1}AS^{-1}\nabla H_q.$$

Unless $\mathcal{H}$ is quadratic, these approximations are different from the ones obtained using the "exact" $H := \mathcal{H}(u_i f_i)$.

A similar approach works for arbitrary $\mathcal{H}$. For $\mathcal{H} = \int h(u_x)$ (and natural boundary conditions), for example, we approximate u_x by $D\mathbf{u}$ and use

$$H_q = \mathbf{w}^{\mathrm{T}}h(D\mathbf{u}), \quad W^{-1}\nabla H_q \;=\; W^{-1}D^{\mathrm{T}}Wh'(D\mathbf{u}) \;=\; W^{-1}A^{\mathrm{T}}h'(D\mathbf{u}).$$

Since $A^{\mathrm{T}}\mathbf{v} = -A\mathbf{v}$ when natural boundary conditions are satisfied, this does then approximate $\frac{\delta H}{\delta u} = -\partial_x h'(u_x)$.

8.3. **Approximating the inner product matrix by quadrature.** Since $\frac{\delta\mathcal{H}}{\delta u} \approx W^{-1}\nabla H_q$, one approach for constructing skew gradient discretizations is to use quadrature in S as well, that is, replace S_{jk} by

$$(S_q)_{jk} = \sum f_j(x_i) f_k(x_i) w_i,$$

so that $S_q = W$ in a cardinal basis. This does yield skew gradient systems $\dot{\mathbf{u}} = W^{-1}AW^{-1}\nabla H_q$. However, $W^{-1}A$ is not an accurate approximation of $\mathcal{D}$: the problem has merely been transfered from $\mathcal{H}$ to $\mathcal{D}$. Applying quadrature to A yields a matrix A_q which is not antisymmetric in general.

8.4. **Summary.** With arbitrary $\mathfrak{F}$, and quadrature approximations of arbitrary Hamiltonians, only very simple operators $\mathcal{D}$ can lead to skew-gradient systems. Suppose $\mathcal{D}$ has a discrete approximation D. Then we get the discretization

$$\dot{\mathbf{u}} = DW^{-1}\nabla H_q,$$

which is a skew-gradient system if DW^{-1} is antisymmetric. The only general cases we know of in which this holds are

(i) On a constant periodic grid, W is a multiple of the identity, so if D is antisymmetric, DW^{-1} is too. This includes the finite difference and Fourier pseudospectral methods discussed in [20].
(ii) On an arbitrary grid, if D is a multiple of the identity, then DW^{-1} is diagonal. In several variables, if each *block* of D is a multiple of the identity, then so is the corresponding block of DW^{-1}. Consequently, if if D is canonical, that is, if $D = \left(\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right)$, then DW^{-1} is antisymmetric if and only if coordinate/momentum pairs use the same quadrature weight.
(iii) On a Legendre zeros grid, $S = W$, and $DW^{-1} = W^{-1}AW^{-1}$ is antisymmetric (discussed above).

The required compatibility between D and W remains an intriguing and frustrating obstacle to the systematic construction of computationally economical conservative discretizations of strongly nonlinear PDEs.

9. Conclusion

Besides discretizations of $\mathcal{D}\frac{\delta\mathcal{H}}{\delta u}$, the method of dual compositions yields new discretizations of linear differential operators $\mathcal{D}$—by taking $\mathcal{H} = \frac{1}{2}\langle u, u\rangle$—and variational derivatives—$\mathcal{D} = 1$. There are obvious generalizations to formally *self*-adjoint $\mathcal{D}$'s and to mixed (e.g., advection-diffusion) operators, where preserving symmetry gives control of the energy. The implementation and accuracy of dual composition discretizations is as for standard "one-step" pseudospectral and Galerkin discretizations, and the three methods coincide in simple cases. Thus, dual composition method discretizations are already in wide usage.

The method as it stands requires that the boundary terms related to $\frac{\delta\mathcal{H}}{\delta u}$ and $\mathcal{D}$ vanish. More general boundary conditions need to be investigated.

Finally, the use of the factorizations of pseudospectral "differentiation" matrices to study stability with the help of relatives of the Kreiss matrix theorem—in particular those found in [14, 18]—deserves a closer look.

Acknowledgements

The authors thank Stanly Steinberg, Rick Beatson, Marc Spijker, Bruce van Brunt, Tom Robey and Grant Lythe for useful conversations. This work was primarily funded by a grant from the Marsden Fund of the Royal Society of New Zealand. R. I. Mclachlan thanks the Mathematical Sciences Research Institute for its hospitality while working on this paper (MSRI acknowledges support through NSF grant no. DMS-9701755). N. Robidoux thanks NAG Inc. for supporting his use of the computer algebra system AXIOM.

References

[1] B. Alpert and V. Rokhlin. A fast algorithm for evaluation of Legendre expansions. *SIAM J. Sci. Comput.*, 12:158–179, 1991.

[2] V. I. Arnold and B. A. Khesin. *Topological Methods in Hydrodynamics.* Springer-Verlag Berlin, 1998.

[3] J. P. Boyd. *Chebyshev & Fourier Spectral Methods*, volume 49 of *Lecture Notes in Engineering.* Springer-Verlag Berlin, Heidelberg, 1989.

[4] Z. Cai, R. Lazarov, T. A. Manteuffel, and S. F. McCormick. First-order system least squares for second-order partial differential equations: part I. *SIAM J. Numer. Anal.*, 31(6):1785–1799, 1994.

[5] M. H. Carpenter and D. Gottlieb. Spectral methods on arbitrary grids. *J. Comput. Phys.*, 129:74–86, 1996.

[6] M. H. Carpenter, D. Gottlieb, and S. Abarbanel. Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: methodology and application to high-order compact schemes. *J. Comput. Phys.*, 111:220–236, 1994.

[7] M. H. Carpenter and J. Otto. High-order “cyclo-difference” techniques: An alternative to finite differences. *J. Comput. Phys.*, 118:242–260, 1995.

[8] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex fourier series. *Math. Comp.*, 19:297–301, 1965.

[9] W. S. Don and D. Gottlieb. The Chebyshev-Legendre method: implementing Legendre methods on Chebyshev points. *SIAM J. Numer. Anal.*, 31(6):1519–1534, 1994.

[10] B. A. Finlayson. *Method of Weighted Residuals and Variational Principles.* Academic Press, New York, 1972.

[11] B. Fornberg. *A Practical Guide to Pseudospectral Methods.* Cambridge University Press, 1996.

[12] I. M. Gelfand and S. V. Fomin. *Calculus of Variations.* Selected Russian Publications in the Mathematical Sciences. Prentice-Hall, Englewood Cliffs, 1963.

[13] D. Gottlieb and L. Lustman. The spectrum of the Chebyshev collocation operator for the heat equation. *SIAM J. Numer. Anal.*, 20:909–921, 1983.

[14] D. Gottlieb and S. A. Orszag. *Numerical analysis of spectral methods: theory and applications.* Number 26 in CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, Philadelphia, 1977.

[15] A. Harten, P. D. Lax, and B. van Leer. On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM Rev.*, 25(1):35–61, 1983.

[16] H. O. Kreiss. Über die stabilitätsdefinition für differenzengleichungen die partielle differentialgleichungen approximieren. *BIT*, 2:153–181, 1962. German with English summary.

[17] H. O. Kreiss and G. Scherer. Finite element and finite difference methods for hyperbolic partial differential equations. In C. de Boor, editor, *Mathematical Aspects of Finite Elements in Partial Differential Equations: Symposium 1974*, pages 195–212. Academic Press, 1974.

[18] D. Levy and E. Tadmor. From semidiscrete to fully discrete: stability of Runge-Kutta schemes by the energy method. *SIAM Rev.*, 40(1):40–73, March 1998.

[19] R. I. McLachlan. Spatial discretization of partial differential equations with integrals. *J. Comput. Phys.* submitted.

[20] R. I. McLachlan. Symplectic integration of Hamiltonian wave equations. *Numer. Math.*, 66:465–492, 1994.

[21] R. I. McLachlan, G.R.W. Quispel, and N. Robidoux. Unified approach to Hamiltonian systems, Poisson systems, gradient systems, and systems with Lyapunov functions or first integrals. *Phys. Rev. Lett.*, 81(12):2399–2403, 1998.

[22] R. I. McLachlan, G.R.W. Quispel, and N. Robidoux. Geometric integration using discrete gradients. *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.*, 357:1021–1046, 1999.

[23] R. I. McLachlan and N. Robidoux. Skew-adjoint finite difference methods. in preparation.

[24] P. Olsson. Summation by parts, projections, and stability. I. *Math. Comput.*, 64:1035–1065, S23–S26, 1995.

[25] P. J. Olver. *Applications of Lie Groups to Differential Equations*, volume 107 of *Graduate Texts in Mathematics*. Springer-Verlag New York, New York, second edition, 1993.

[26] L. G. Reyna. l_2 estimates for Chebyshev collocation. *J. Sci. Comput.*, 3(1):1–23, 1988.

[27] M. Shashkov. *Conservative Finite-Difference Methods on General Grids*. CRC Press, Boca Raton, 1996.

[28] M.N. Spijker and F.A.J. Straetemans. Stability estimates for families of matrices of nonuniformly bounded order. *Linear Algebra Appl.*, 239:77–102, 1996.

[29] B. Strand. Summation by parts for finite difference approximations for d/dx. *J. Comput. Phys.*, 110:47–67, 1994.

[30] J. C. Strickwerda and B. A. Wade. A survey of the Kreiss matrix theorem for power bounded families of matrices and its extensions. In *Linear Operators (Warsaw, 1994)*, number 38 in Banach Center Publ., pages 339–360, Warsaw, 1997. Polish Acad. Sci.

[31] J. C. Strikwerda. *Finite difference schemes and partial differential equations*. Advanced Books & Software. Wadsworth & Brooks/Cole, Pacific Grove, California, 1989.

[32] L. N. Trefethen. Pseudospectra of linear operators. In K. Kirchgassner, O. Mahrenholtz, and R. Mennicken, editors, *ICIAM 95*, pages 401–434, Berlin. Akademie Verlag. Reprinted as [33].

[33] L. N. Trefethen. Pseudospectra of linear operators. *SIAM Rev.*, 39(3):383–406, 1977. Previously published as [32].

[34] L. N. Trefethen. Lax-stability vs. eigenvalue stability of spectral methods. In K.W. Morton and M.J. Baines, editors, *Numerical Methods for Fluid Dynamics III*, Oxford, 1988. Clarendon Press.

[35] L. N. Trefethen. Finite difference and spectral methods for ordinary and partial differential equations. Book draft, available from `http://www.comlab.ox.ac.uk/oucl/users/nick.trefethen/pdetext.html`, 1989.

[36] L. N. Trefethen. Pseudospectra of matrices. In D.F. Griffiths and G.A. Watson, editors, *Numerical Analysis 1991*, pages 234–266. Longman Sci. Tech. Publ., 1992.

[37] L. N. Trefethen and M.R. Trummer. An instability phenomenon in spectral methods. *SIAM J. Numer. Anal.*, 24:1008–1023, 1987.

[38] J. L. M. van Dorsselaer, J. F. B. M. Kraaijevanger, and M. N. Spijker. Linear stability analysis in the numerical solution of initial value problems. *Acta Numerica*, 2:199–237, 1993.

[39] B. A. Wade. Symmetrizable finite difference operators. *Math. Comp.*, 54(190):525–543, April 1990.

[40] L. B. Wahlbin. *Superconvergence in Galerkin finite element methods*, volume 1605 of *Lecture Notes in Mathematics*. Springer-Verlag New York, 1995.

[41] J. A. C. Weideman and L. N. Trefethen. The eigenvalues of second-order spectral differentiation matrices. *SIAM J. Numer. Anal.*, 25:1279–1298, 1988.

[42] D. Zwillinger. *Handbook of Differential Equations*. Academic Press, San Diego, second edition, 1989.

Mathematics, Institute of Fundamental Sciences, Massey University, Private Bag 11–222, Palmerston North, New Zealand 5301.

E-mail address: `R.McLachlan@massey.ac.nz, N.Robidoux@massey.ac.nz`.