JOURNAL OF COMPUTATIONAL PHYSICS **129,** 121–133 (1996)
ARTICLE NO. 0238

On Numerical Methods for Hamiltonian PDEs and a Collocation Method for the Vlasov–Maxwell Equations

James Paul Holloway*

Department of Nuclear Engineering, Cooley Building, North Campus, University of Michigan, Ann Arbor, Michigan 48109-2104

Received April 20, 1995; revised February 9, 1996

Hamiltonian partial differential equations often have implicit conservation laws—constants of the motion—embedded within them. It is not, in general, possible to preserve these conservation laws simply by discretization in conservative form because there is frequently only one explicit conservation law. However, by using weighted residual methods and exploiting the Hamiltonian structure of the equations it is shown that at least some of the conservation laws are preserved in a method of lines (continuous in time). In particular, the Hamiltonian can always be exactly preserved as a constant of the motion. Other conservation laws, in particular linear and quadratic Casimirs and momenta, can sometimes be conserved too, depending on the details of the equations under consideration and the form of discretization employed. Collocation methods also offer automatic conservation of linear and quadratic Casimirs. Some standard discretization methods, when applied to Hamiltonian problems are shown to be derived from a numerical approximation to the exact Poisson bracket of the system. A method for the Vlasov–Maxwell equations based on Legendre–Gauss–Lobatto collocation is presented as an example of these ideas.

1. INTRODUCTION

Some partial differential and integro-partial differential equations of evolution have an underlying Hamiltonian structure [1–4] that is responsible for the existence of interesting conservation laws. Since the Hamiltonian structure of these equations is tied closely to the physical system that they represent it would seem prudent to retain as much of this structure as possible when developing numerical methods to discretize such problems. Ideally, we should like to be able to develop semi-discrete methods—continuous in time—that are finite dimensional Hamiltonian systems; these might then be fully discretized with either symplectic or energy conserving time discretization methods [5–8]. In this paper the less ambitious goal of preserving some of the constants of the motion from the continuous problem will be pursued; for example, I will show that a properly formulated weighted residuals method can always exactly conserve the Hamiltonian, even at finite spatial resolution, no matter how nonlinear a functional the Hamiltonian is.

The trivial prototype for a Hamiltonian partial differential equation is the linear advection equation

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x}. \tag{1}$$

This equation is, of course, in the form of a classic hyperbolic conservation law; thus, the linear advection equation is simply a statement of the conservation of "mass" $\int u\,dx$, so that

$$\frac{d}{dt}\int_{-\infty}^{\infty} u\,dx = 0. \tag{2}$$

But, assuming that everything decays quickly as $|x| \to \infty$, any classical solution of Eq. (1) also conserves "energy" $\int u^2\,dx$, and indeed, it conserves $\int u^n\,dx$ for any n; that is,

$$\frac{d}{dt}\int_{-\infty}^{\infty} u^n\,dx = 0. \tag{3}$$

These constants of motion can be identified with algebraic features of the Hamiltonian structure that underlies the linear advection problem; the philosophy presented in this paper is to exploit this structure as much as possible in deriving (continuous in time) numerical methods to ensure that the numerical approximation has something like Hamiltonian structure and to ensure that this structure is closely enough related to its continuous cousin that it preserves some of the constants of the motion.

It is important to note that the conservation laws in Eq. (3) for $n > 1$ are not explicit; Eq. (1) is a hyperbolic conservation law for conservation of mass, but we do not have separate hyperbolic conservation laws for conservation of "energy" $\int u^2\,dx$ or for the conservation of $\int u^3\,dx$, whatever that is. This is typical of Hamiltonian partial differential equations; unlike, for example, the Navier–Stokes equations which include explicit equations

* Supported by NSF Grants CCR-9110352 and ECS-9358344. E-mail: hagar@umich.edu.

0021-9991/96 $18.00

for conservation of mass, momentum, and energy, Hamiltonian systems often have conserved quantities—also called constants of motion—that are direct consequences of the governing equation, but not expressed as separate conservation laws. For example, the conservation of the quantities $\int u^n\,dx$ for $n > 1$ is a consequence of Eq. (1), but their conservation cannot be guaranteed simply by discretizing the equation in conservative form. For example, the upwind difference on a uniform grid of unit spacing,

$$\frac{du_i}{dt} = u_{i+1} - u_i, \quad i = -\infty \cdots \infty, \tag{4}$$

conserves $\sum_i u_i$, but not $\sum_i u_i^2$; similarly, the central difference conserves both these sums, but not $\sum_i u_i^3$. The advantage of the central difference, in this respect, comes about because it happens to retain more of the Hamiltonian structure than did the forward difference.

In the abstract, Hamiltonian systems are basically built out of two things: a functional of the system state called the Hamiltonian—often but not necessarily the physical energy in the state—and a bilinear operator called a Poisson bracket that can build a new functional of system state from two old ones. The key point of this paper is that we should focus on the bracket and discretize it, producing thereby a ''numerical bracket'' which can be used to build a Hamiltonian-like system on a finite dimensional space of numerical degrees-of-freedom. By examining the properties of this discretization of the bracket we can determine beforehand if certain conservation laws will be preserved, in much the same way that discretization of hyperbolic conservation laws in conservative form preserves their conservative structure.

The purpose of this paper is therefore to examine how the conserved quantities—and particularly the Hamiltonian itself—of Hamiltonian partial differential equations can be preserved in weighted residual and collocation methods by examining the properties of the Poisson bracket. I believe that this study provides some insight into why some methods simultaneously preserve several of the constants of the motion of Hamiltonian PDEs while others do not, and it thereby provides us with a set of ideas on how to develop conservative methods for Hamiltonian problems—what basis functions or degrees of freedom to choose—or how to recognize the conservative Hamiltonian structure in standard methods.

2. BRACKETS

The methods to be presented in this paper are based on purposefully exploiting the structure of Hamiltonian PDEs as *Hamiltonian* PDEs, rather than exploiting some other feature such as their structure as hyperbolic conservation laws. A very brief review of the pertinent Hamiltonian structure for autonomous isolated systems is presented here. All of the material in this section can be found in standard texts on Hamiltonian systems (see, for example, Goldstein [9], Olver [10], and Marsden and Ratiu [11]).

The problems of interest are evolution equations for a system state vector $u(t)$, which can represent several functions of several independent variables besides time. This state is assumed to satisfy an evolution equation of the form

$$\frac{\partial u}{\partial t} = \mathscr{J}(u)\frac{\delta \mathscr{H}}{\delta u}(u), \tag{5}$$

where $\mathscr{J}(u)$, for a fixed u, is a linear operator that is antisymmetric with respect to an inner product $\langle\ ,\ \rangle$, and $\delta\mathscr{H}/\delta u$ is the variational derivative of a real valued nonlinear functional $\mathscr{H}$, called the Hamiltonian of the system. This derivative is the function $\delta\mathscr{H}/\delta u$ that satisfies

$$\lim_{\varepsilon\to 0}\frac{\mathscr{H}(u+\varepsilon h) - \mathscr{H}(u)}{\varepsilon} = \left\langle \frac{\delta \mathscr{H}}{\delta u}, h\right\rangle \tag{6}$$

for each arbitrary perturbation h. This essentially plays the role of gradient of the Hamiltonian functional $\mathscr{H}$. It can generally depend explicitly on u and on all of the independent variables on which u depends, excepting only time (this is the autonomous case; many of the ideas presented in this paper also apply to the nonautonomous case, but this is not further discussed). For example, Eq. (1) (with $u(x, t) \to 0$ as $|x| \to \infty$) is in the form of Eq. (5), with $\mathscr{J} = \partial/\partial x$ and

$$\mathscr{H}(u) = \frac{1}{2}\int_{-\infty}^{\infty} u^2\,dx \tag{7}$$

which has $\delta\mathscr{H}/\delta u = u$.

Now, suppose that $u(t)$ represents a solution of Eq. (5), and let $\mathscr{F}$ be any (generally nonlinear) functional of u; then we easily compute that

$$\begin{aligned}\frac{d}{dt}\mathscr{F}(u(t)) &= \left\langle \frac{\delta \mathscr{F}}{\delta u}(u(t)), \frac{du}{dt}\right\rangle \\ &= \left\langle \frac{\delta \mathscr{F}}{\delta u}(u(t)), \mathscr{J}(u(t))\frac{\delta \mathscr{H}}{\delta u}(u(t))\right\rangle.\end{aligned} \tag{8}$$

So consider the creature $\{\{\ ,\ \}\}$, called the bracket, which takes any two functionals, $\mathscr{F}$ and $\mathscr{G}$, and builds a new one, $\{\{\mathscr{F}, \mathscr{G}\}\}$ according to the formula

$$\{\{\mathscr{F}, \mathscr{G}\}\}(u) = \left\langle \frac{\delta \mathscr{F}}{\delta u}(u), \mathscr{J}(u)\frac{\delta \mathscr{G}}{\delta u}(u)\right\rangle. \tag{9}$$

Equation (8) says that, in terms of this bracket, the rate of change of a functional $\mathscr{F}$ along a solution of Eq. (5) is given by

$$\frac{d}{dt}\mathscr{F}(u(t)) = \{\{\mathscr{F}, \mathscr{H}\}\}(u(t)). \tag{10}$$

Because of the properties of $\mathscr{I}(u)$, and of the variational derivative, the bracket has the following three properties:

$$\{\{\alpha\mathscr{F} + \beta\mathscr{G}, \mathscr{H}\}\} = \alpha\{\{\mathscr{F}, \mathscr{H}\}\} + \beta\{\{\mathscr{G}, \mathscr{H}\}\} \tag{11}$$

$$\{\{\mathscr{F}, \mathscr{G}\}\} = -\{\{\mathscr{G}, \mathscr{F}\}\} \tag{12}$$

$$\{\{\mathscr{F}\mathscr{G}, \mathscr{H}\}\} = \mathscr{F}\{\{\mathscr{G}, \mathscr{H}\}\} + \mathscr{G}\{\{\mathscr{F}, \mathscr{H}\}\}. \tag{13}$$

Thus, this bracket is bilinear (Eq. (11)), it is anti-symmetric (Eq. (12)), and it satisfies the chain rule (Eq. (13)). Each of these properties is inherited: bilinearity from the inner product, anti-symmetry from $\mathscr{I}$, and the chain rule from the variational derivative. If, in addition, the bracket satisfies the rather nontrivial Jacobi condition

$$\{\{\{\mathscr{F}, \mathscr{G}\}\}, \mathscr{H}\}\} + \{\{\{\mathscr{G}, \mathscr{H}\}\}, \mathscr{F}\}\} + \{\{\{\mathscr{H}, \mathscr{F}\}\}, \mathscr{G}\}\} = 0 \tag{14}$$

for all functionals $\mathscr{F}$, $\mathscr{G}$, $\mathscr{H}$ then the bracket is called a Poisson bracket. In this paper I shall be concerned with problems that can be written in the form of Eq. (5) and for which $\mathscr{I}$ generates a bracket that has the properties of Eqs. (11)–(13), but not necessarily Eq. (14). Authentic Hamiltonian systems are based on true Poisson brackets which must satisfy Eq. (14). But while the Jacobi condition is responsible for much of the analytic beauty of Hamiltonian systems, the Jacobi condition is not crucial to the existence of constants of the motion, and it is at times not satisfied by the brackets presented in this paper. In this paper the goal is conservative evolution, so the Jacobi condition will not be needed.

For a given Hamiltonian $\mathscr{H}$, any other functional $\mathscr{F}$ that satisfies $\{\{\mathscr{F}, \mathscr{H}\}\} = 0$ is obviously a constant of motion defined by that Hamiltonian; for, as we see from Eq. (10), the quantity $\mathscr{F}(u(t))$ then equals $\mathscr{F}(u(0))$. For evolution systems derived from brackets these constants can be classified into three types: the Hamiltonian, Casimirs, and momenta. The Hamiltonian $\mathscr{H}$ has a special place among the constants of the motion because its gradient defines the evolution of the system; it is constant simply because the bracket is anti-symmetric: $\{\{\mathscr{H}, \mathscr{H}\}\} = -\{\{\mathscr{H}, \mathscr{H}\}\} = 0$. On the other hand, with Hamiltonian PDEs it is common for the bracket to be degenerate, so that there is a functional $\mathscr{C}$ that satisfies $\{\{\mathscr{C}, \mathscr{G}\}\} = 0$ for all other functionals $\mathscr{G}$. In this case $\mathscr{C}$ is conserved no matter what the Hamiltonian is; such constants are called Casimirs. Finally, other constants of motion $\mathscr{M}$ that are neither the Hamiltonian nor a Casimir are generally called momenta. There is a known correspondence between symmetries of the Hamiltonian and the momenta (see, e.g., [9–11]); the classical example of this is the linear momentum of a particle in a mechanical system, which is conserved provided only that the Hamiltonian for the system is invariant under the group of translations in space.

The simplest examples of Hamiltonian partial differential equations are based on the bracket

$$\{\{\mathscr{F}, \mathscr{G}\}\} = \int_{-\infty}^{\infty} \frac{\delta\mathscr{F}}{\delta u}\frac{\partial}{\partial x}\frac{\delta\mathscr{G}}{\delta u}\,dx \tag{15}$$

(assuming rapid decay as $|x| \to \infty$). The functional $\mathscr{H}_{\mathrm{LA}}(u) = \int u^2/2\,dx$, with $\delta\mathscr{H}_{\mathrm{LA}}/\delta u = u$, is then the Hamiltonian for the linear advection dynamics

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\frac{\delta\mathscr{H}}{\delta u} = \frac{\partial}{\partial x}u. \tag{16}$$

Similarly the Hamiltonian $\mathscr{H}_{\mathrm{NA}}(u) = \int u^3/6\,dx$, with $\delta\mathscr{H}_{\mathrm{NA}}/\delta u = u^2/2$, generates the nonlinear advection dynamics

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\frac{\delta\mathscr{H}}{\delta u} = \frac{\partial}{\partial x}\frac{u^2}{2} = u\frac{\partial u}{\partial x} \tag{17}$$

while this same bracket structure and the Hamiltonian

$$\mathscr{H}_{\mathrm{KdV}}(u) = \int_{-\infty}^{\infty}\left[\frac{1}{6}u^3 - \frac{1}{2}\left(\frac{\partial u}{\partial x}\right)^2\right]dx \tag{18}$$

produce the dynamics

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\frac{\delta\mathscr{H}}{\delta u} = \frac{\partial}{\partial x}\left(\frac{u^2}{2} + \frac{\partial^2 u}{\partial x^2}\right) = u\frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3} \tag{19}$$

which is the Korteweg–de Vries (KdV) equation [12].

These three systems–linear advection, nonlinear advection, and the Korteweg–de Vries equation—can all be written as hyperbolic conservation laws, and thus all three conserve the "mass" functional $\mathscr{N}(u) = \int_{-\infty}^{\infty} u\,dx$. From the Hamiltonian systems perspective we see that $\delta\mathscr{N}/\delta u = 1$, so for any functional $\mathscr{G}$

$$\{\{\mathscr{N}, \mathscr{G}\}\} = -\{\{\mathscr{G}, \mathscr{N}\}\} = -\int\frac{\delta\mathscr{G}}{\delta u}\frac{\partial}{\partial x}\frac{\delta\mathscr{N}}{\delta u}\,dx = 0. \tag{20}$$

Thus, $\mathscr{N}$ is a Casimir; it is conserved no matter what is used for the Hamiltonian. However, a direct computation will show that $\mathscr{H}_{\mathrm{LA}}$ is a momentum for both the nonlinear

advection equation and the KdV equation. Note that $\int u^2/2\, dx$ is often called "energy" because it is a quadratic quantity, like the Hamiltonian of a harmonic oscillator, but formally it is a momentum, a quantity conserved because of a symmetry of the Hamiltonian, in this case invariance under spatial translations. Thus, the structure of these equations as hyperbolic conservation laws tells us only about a single conserved quantity, the Casimir $\mathscr{N}$, but in fact they conserve quantities other than this as well; and while discretizing them in conservative form will ensure that $\mathscr{N}$ is conserved, doing so does not guarantee the conservation of their respective Hamiltonians, or of the momentum $\mathscr{H}_{\mathrm{LA}}$.

The theme of this paper is that it is possible, in general, to discretize Hamiltonian partial differential equations so that the resulting numerical system can, in the limit of continuous time, be derived from a bracket. Since it is this bracket structure that is responsible for constants of the motion in Hamiltonian systems, this view of discretization provides a great deal of insight into how to discretize so that conserved quantities are retained in the numerical method.

3. THE NUMERICAL BRACKET FOR A WEIGHTED RESIDUALS METHOD

Suppose that we wish to develop a weighted residuals method for the Hamiltonian equation

$$\frac{\partial u}{\partial t} = \mathscr{I}(u)\frac{\delta \mathscr{H}}{\delta u}(u) \tag{21}$$

based on some finite expansion of the sort $u \approx U = a^i \phi_i$ (the summation convention is in effect, with summation implied over any index repeated in both super- and subscript). Here a^i is a time-dependent expansion coefficient and ϕ_i is an expansion function which depends on the independent variables—other than time—of the problem. We can, of course, view this problem as one of finding a set of ODEs for the expansion coefficients a^i. In a weighted residuals method each of the expansion coefficients is ideally a linear functional of u, $a^i = \langle \phi^i, u\rangle$, where ϕ^i is a weight function for the method. But in a Hamiltonian problem the time rate of change of functionals is determined by the bracket; one approach to the problem, and the one put forward here, is therefore to simply modify the exact bracket $\{\{\ ,\ \}\}$ so that it acts not on functionals of the continuous state u, but on the discrete state represented by the expansion coefficients a^i.

So let us begin with a set of N expansion functions $\{\phi_i\}_{i=1\cdots N}$ and a set of N weight functions $\{\phi^i\}_{i=1\cdots N}$, which are required to satisfy the bi-orthonormality conditions

$$\langle \phi^i, \phi_j\rangle = \delta^i_j. \tag{22}$$

This can always be arranged, of course, through a Gram–Schmidt process. From these functions we can construct a projection operator P,

$$Pu = \langle \phi^i, u\rangle \phi_i, \tag{23}$$

and its adjoint $P^\dagger$,

$$P^\dagger u = \langle \phi_i, u\rangle \phi^i. \tag{24}$$

A standard (continuous in time) weighted residuals method for Eq. (5) is then formulated by projecting the initial data onto the expansion functions $U(0) = Pu(0)$ and similarly projecting Eq. (5) to get

$$\frac{\partial U}{\partial t} = P\mathscr{I}(U)\frac{\delta \mathscr{H}}{\delta u}(U). \tag{25}$$

Unfortunately, this system of evolution equations cannot, in general, be derived from a bracket, so the Hamiltonian $\mathscr{H}$ is not, in general, conserved by this weighted residuals discretization.

The heart of this paper is to point out that a trivial modification of Eq. (25) produced by placing the adjoint projection operator $P^\dagger$ on the right of the variational derivative of $\mathscr{H}$ in Eq. (25),

$$\frac{\partial U}{\partial t} = P\mathscr{I}(U)P^\dagger \frac{\delta \mathscr{H}}{\delta u}(U), \tag{26}$$

is in fact derived from the bracket

$$[\mathscr{F}, \mathscr{G}]_w = \left\langle \frac{\delta \mathscr{F}}{\delta u}, P\mathscr{I}(u)P^\dagger \frac{\delta \mathscr{G}}{\delta u}\right\rangle \tag{27}$$

(the subscript "w" denoting weighted residuals). This numerical bracket is clearly bilinear and satisfies the chain rule, and, because $\mathscr{I}$ is sandwiched between P and $P^\dagger$, this bracket also retains the anti-symmetry of $\mathscr{I}$. While this bracket will not necessarily satisfy the Jacobi identity, it does contain all the properties of a bracket, Eqs. (11)–(13), that are useful in seeking constants of the motion.

The philosophical difference between Eq. (25) and Eq. (26) is this: in a standard weighted residuals method we project the rate of change of u onto the subspace that is spanned by the expansion functions, while in Eq. (26) we replace $\mathscr{I}$ by $P\mathscr{I}P^\dagger$ in order to retain the anti-symmetry of the operator $\mathscr{I}$ in the numerical method. Both approaches are intended to derive an approximate description of the dynamics that leave a subspace invariant, but in Eq. (26) we may (although not necessarily) accept some extra error (due to $P^\dagger$) in order to produce a numerical

description that can be derived from a bracket, and thereby achieve conservation of the Hamiltonian.

Because Eq. (26) inherits its bracket structure from the bracket of the original system, Eq. (5), it is easy to make some observations about its conservation properties, i.e., to note how it can preserve some constants of the motion. The most trivial, yet most important, of these conservation properties (CP) is:

CP 3.1. *Equation* (26) *always conserves the Hamiltonian* $\mathscr{H}$.

Proof. This follows trivially from the fact that the numerical bracket $[\ ,\ ]_w$ is anti-symmetric. ▮

This means that, in the limit of continuous time but otherwise finite discretization, the Hamiltonian can be exactly conserved, no matter how nasty a nonlinear functional it may be.

The next two conservation properties deal with two special cases in which Casimirs are relatively easy to retain:

CP 3.2. *A linear Casimir* $\mathscr{C}(u) = \langle c, u\rangle$ *of a bracket* $\{\{\ ,\ \}\}$ *is also a Casimir of the numerical bracket* $[\ ,\ ]_w$ *if* $P^\dagger c = c$.

Proof. Such a Casimir obviously satisfies $P\mathscr{J}(u)P^\dagger \delta\mathscr{C}/\delta u = P\mathscr{J}(u)c$. But since $\mathscr{C}$ is a Casimir of the exact bracket $\{\{\ ,\ \}\}$ it must be that $\mathscr{J}(u)c = 0$, and so $[\mathscr{G}, \mathscr{C}]_w = 0$ for any functional $\mathscr{G}$. ▮

CP 3.3. *A quadratic Casimir* $\mathscr{C}(u)$ *with* $\delta\mathscr{C}/\delta u = u$ *is constant along solutions of Eq.* (26) *if* $P = P^\dagger$.

Proof. In general,

$$\frac{d}{dt}\mathscr{C}(U(t)) = -[\mathscr{H}, \mathscr{C}]_w(U(t))$$

and in the present case $P\mathscr{J}(U)P^\dagger \delta C/\delta u = P\mathscr{J}(U)PU$. But $PU = U$ since P is a projection, so we have $P\mathscr{J}(U)P^\dagger\delta\mathscr{C}/\delta u = P\mathscr{J}(U)U = 0$ since $\mathscr{C}$ is a Casimir of the exact bracket. Hence $[H, C]_w(U(t)) = 0$, which implies $dC/dt = 0$. ▮

The last two conservation properties provide similar results for momenta:

CP 3.4. *A momentum* $\mathscr{M}(u) = \langle m, u\rangle$ *that is linear in* u *is conserved by Eq.* (26) *if* $P^\dagger m = m$ *and* $P\mathscr{J}((U)m = \mathscr{J}(U)m$ *for all* U *in the range of* P.

Proof. In this case $P\mathscr{J}(U))P^\dagger m = \mathscr{J}(U)m$, and so

$$\frac{d\mathscr{M}}{dt} = \left\langle m, P\mathscr{J}P^\dagger\frac{\delta\mathscr{H}}{\delta u}\right\rangle = -\left\langle\frac{\delta\mathscr{H}}{\delta u}, P\mathscr{J}P^\dagger m\right\rangle$$
$$= -\left\langle\frac{\delta\mathscr{H}}{\delta u}, \mathscr{J}m\right\rangle = 0$$

since $\mathscr{M}$ is a momentum of the continuous system. ▮

CP 3.5. *A quadratic momentum with* $\delta\mathscr{M}/\delta u = u$ *is conserved by Eq.* (26) *if* $P = P^\dagger$ *and* $P\mathscr{J}(U)U = \mathscr{J}(U)U$ *for all* U *in the range of* P.

Proof. Now we have $P\mathscr{J}(U)P^\dagger\delta\mathscr{M}/\delta u = P\mathscr{J}(U)P\delta\mathscr{M}/\delta u = P\mathscr{J}(U)U = \mathscr{J}(U)U$, for U in the range of P. And so $d\mathscr{M}/dt = \langle U, P\mathscr{J}P^\dagger\delta\mathscr{H}/\delta u\rangle = -\langle\delta\mathscr{H}/\delta, \mathscr{J}U\rangle = 0$. ▮

So it is possible in general to write a weighted residuals type method in such a way that at least the Hamiltonian and possibly several other constants of the motion are preserved, and this is true independent of the details of the Hamiltonian of the system being treated. Sometimes these methods are standard weighted residuals methods; when $P\mathscr{J}(U) = P\mathscr{J}(U))P^\dagger$ Eqs. (26) and (25) are identical, otherwise they are not. The conservation of quadratic momenta and Casimirs (CP 3.3 and CP 3.5) have a common requirement that the projection be self-adjoint $P = P\dagger$; this is reminiscent of a known result for weighted residuals methods [13].

It is standard in finite dimensional Hamiltonian mechanics to write a bracket in terms of its structure functions, which are the brackets between all pairs of degrees of freedom. For the expansion coefficients a^i these structure functions are denoted $[a^i, a^j]_w$. This function relates the time rate-of-change of the dynamical variable a_i to the rate of change of the Hamiltonian with respect to a_j, and normally it is a function of the system state. Using the fact that $P^\dagger$ is a projection, these structure functions can be evaluated from

$$[a^i, a^j]_w(\mathbf{a}) = \langle\phi^j, \mathscr{J}(a^k\phi_k)\phi^j\rangle, \tag{28}$$

where $\mathbf{a} = \{a^1, a^2, ..., a^N\}$ is the vector of expansion coefficients. Note that these are in general functions of the expansion coefficients $\mathbf{a}$ because $\mathscr{J}$ is such a function. The utility of the structure functions is that they allow Eq. (26) to be rewritten as a system of ODEs in slightly different form. We start from Eq. (26) in the form

$$\frac{da^i}{dt} = \left\langle\phi^j, P\mathscr{J}(a^k\phi_k)P^\dagger\frac{\delta\mathscr{H}}{\delta u}(a^k\phi_k)\right\rangle, \tag{29}$$

where U has been rewritten as the explicit expansion $a^k\phi_k$ (and remember the summation convention). But from the definition of the variational derivative

$$\left\langle\phi_j, \frac{\delta\mathscr{H}}{\delta u}(a^k\phi_k)\right\rangle = \lim_{\varepsilon\to 0}\frac{\mathscr{H}(a^k\phi_k + \varepsilon\phi_j) - \mathscr{H}(a^k\phi_k)}{\varepsilon}$$
$$= \frac{\partial}{\partial a^j}\mathscr{H}(a^k\phi_k) \tag{30}$$

so

$$P^\dagger \frac{\delta \mathscr{H}}{\delta u}(a^k \phi_k) = \frac{\partial}{\partial a^j}\mathscr{H}(a^k\phi_k)\phi^j \tag{31}$$

(remember to sum over j). Using the linearity of $\mathscr{I}$ and noting that $\partial\mathscr{H}/\partial a^j$ are scalars, it then follows that

$$\frac{da^i}{dt} = \langle \phi^i, \mathscr{I}(a^k\phi_k)\phi^j\rangle \frac{\partial}{\partial a^j}\mathscr{H}(a^k\phi_k) \tag{32}$$

or

$$\frac{da^i}{dt} = [a^i, a^j]_w \frac{\partial \mathscr{H}}{\partial a^j}, \tag{33}$$

which is in the form of a classical finite dimensional Hamiltonian system in noncanonical coordinates. While algebraically equivalent to Eq. (26), this form of the equations emphasizes a different view of the weighted residuals method; it forces attention on the bracket between the weight functions, $[a^i, a^j]$, as the quantities that describe a class of numerical methods applicable to a problem derived from a particular bracket.

The developments above emphasize that it is always possible to write down a weighted residuals method in a form that can be derived from a numerical bracket and which, therefore, exactly conserves the Hamiltonian of the exact system. Because the numerical bracket is derived from the true bracket using projection operators it is also fairly straightforward to see how some other conserved quantities—Casimirs and momenta—can sometimes be retained by the weighted residuals method as well.

3.1. *Example: Fourier–Galerkin Methods*

The ideas developed above in the abstract will now be applied to the Fourier–Galerkin method; in this section we shall see the highly conservative nature of Fourier spectral methods as applied to advection problems and show that this conservation comes about because these methods, applied to these problems, in fact, yield finite dimensional Hamiltonian systems of the form given in Eq. (26) or (33).

We begin with the Poisson bracket given in Eq. (15), where $\mathscr{I} = \partial/\partial x$, and use periodic boundary conditions. The obvious set of expansion functions are then the sines and cosines, so let us represent the projections $P = P^\dagger$ as

$$Pu = \frac{1}{2}C^0 + \sum_{n=1}^{N} C^n \cos(nx) + \sum_{n=1}^{N} S^n \sin(nx), \tag{34}$$

where $C^n = (1/\pi)\int_{-\pi}^{\pi} u(x)\cos(nx)\,dx$ and $S^n = (1/\pi)\int_{-\pi}^{\pi} u(x)\sin(nx)\,dx$. From these we can easily compute the structure functions of the numerical bracket using Eq. (28) and discover that

$$\begin{aligned}[C^n, C^m]_w &= [S^n, S^m]_w = 0\\ [C^n, S^m]_w &= -[S^n, C^m]_w = \frac{n}{\pi}\delta^{nm}.\end{aligned} \tag{35}$$

Because these structure functions are independent of C^n and S^n this bracket actually satisfies the Jacobi identity too; indeed, it is almost in canonical form. The evolution equations derived from this bracket and any Hamiltonian $\mathscr{H}$ are then

$$\begin{aligned}\frac{dC^n}{dt} &= \frac{n}{\pi}\frac{\partial \mathscr{H}}{\partial S^n}\\ \frac{dS^n}{dt} &= -\frac{n}{\pi}\frac{\partial \mathscr{H}}{\partial C^n}\end{aligned} \tag{36}$$

which is quite obviously in Hamiltonian form with S^n the variable dynamically conjugate to C^n.

We see immediately that, no matter what the Hamiltonian is, C^0 is a Casimir and is exactly conserved by the numerical approximation; physically it is the mass function $\mathscr{N} = \int u\,dx$, and its conservation was inevitable from CP 3.2. Also, CP 3.5 applies, so that $\mathscr{H}_{\mathrm{LA}} = \int u^2\,dx$ will be conserved by the numerical method if it is conserved by the exact system. And, of course, no matter what the Hamiltonian $\mathscr{H}$ is, it will also be conserved according to CP 3.1. Furthermore, for the projection P and operator $\mathscr{I}$ at hand we have $P\mathscr{I}P^\dagger = P\mathscr{I}$, and the method of Eq. (26) is therefore identical to the standard weighted residuals method (in this case the Fourier–Galerkin spectral method) given in Eq. (25), which therefore has a numerical Poisson bracket hidden behind it.

Thus, without doing any specific checking, we immediately know that a standard Fourier–Galerkin spectral method applied to the linear advection equation, to the nonlinear advection equation, or to the KdV equation will exactly conserve $\int u\,dx$, $\int u^2\,dx$, and their respective Hamiltonians (in the limit of continuous time, of course). Indeed, the conclusion holds for any system derived from a Hamiltonian of the form $\mathscr{H}(u) = \int g(u, u_x, u_{xx}, \ldots)\,dx$ whose integrand is explicitly independent of x (necessary and sufficient for the momentum $\mathscr{H}_{\mathrm{LA}}$ to be conserved); that is, the Fourier–Galerkin spectral method applied to any PDE of the form

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\sum_{n=0}(-1)^n \frac{\partial^n}{\partial x^n} D_n g(u, u_x, u_{xx}, \ldots), \tag{37}$$

where $D_n g$ denotes the derivative of g with respect to its

nth argument, will exactly conserve $\int u(x, t)\, dx$, $\int u^2(x, t)\, dx$, and $\mathscr{H}$ itself.

It is especially interesting to note that the class of methods suggested in this section were actually used by Gardner [12] in his analytic development of the bracket, Eq. (15), and Hamiltonian, Eq. (18), of the KdV equation.

4. BRACKETS FOR COLLOCATION METHODS

Of course, a Fourier–Galerkin spectral method as described in Section 3.2 is not particularly efficient for the solution of nonlinear problems because the evaluation of the nonlinear terms can be rather expensive; instead, Fourier collocation is generally preferred. In a collocation method the continuum of state variables $u(x)$ is replaced by the discrete set of values u^i representing $u(x)$ on a set of grid points x_i. This can be understood as a weighted residuals method using weight functions that are delta functions located at the grid points, $\delta(x - x_i)$, or it can also be understood as a weighted residuals method based on a discrete approximation to the inner product. The former approach is not useful here since typically $\langle \delta(x - x_i), \mathscr{I}\delta(x - x_j)\rangle = 0$ if $i \neq j$ and is often undefined if $i = j$, formally because δ is not in the range of $\mathscr{I}$. Therefore it is the latter approach, approximation of the inner product, that will be exploited here. The key idea will be that a collocation method can be fit within the bracket structure of a problem by replacing the continuum inner product by a discrete quadrature rule and by replacing the operator $\mathscr{I}$ by an approximation J computed using data on the grid, provided only that the approximation to $\mathscr{I}$ retains its antisymmetry under the quadrature formula chosen (which usually means that a discrete integration-by-parts rule must hold).

Suppose that the inner product $\langle\ ,\ \rangle$ can be approximated using a quadrature rule, so that

$$\langle f, g\rangle \approx \sum_k w_k f(x_k) g(x_k), \tag{38}$$

where the w_k represent the quadrature weights on the grid x_k, and note that x_k could represent a point in a one, two, three, or more dimensional set of independent variables. Also note that the quadrature weights will quite naturally allow us to lower and raise indices using the diagonal tensor $w_{ij} = w_i\delta_{ij}$ and its inverse $w^{ij} = (w_i)^{-1}\delta^{ij}$ (no sum). Now let $C_i(x)$ represent a cardinal function for the grid, meaning that $C_i(x_k) = \delta_i^k$, and define $C^i(x) = w^{ij}C_j(x)$. Note that these functions satisfy $C^i(x_k) = w^{ik} = (w_i)^{-1}\delta^{ik}$ so that as the number of mesh points is increased and the quadrature weights go to zero the functions $C^i(x)$ will become more and more delta-like. Together, $C_i(x)$ and $C^j(x)$ provide a bi-orthonormal set of expansion and weight functions under the discrete inner product; that is,

$$\langle C_i, C^j\rangle \approx \sum_k w_k C_i(x_k) C^j(x_k) = \delta_i^j. \tag{39}$$

Letting $C_i^k = C_i(x_k)$ and $C^{ik} = w^{ij}C_j^k = C^i(x_k)$ we can also write this as

$$\langle C_i, C^j\rangle \approx \langle C_i, C^j\rangle_c = C_i^k w_{kl} C^{jl} = \delta_i^j, \tag{40}$$

where $\langle\ ,\ \rangle_c$ represents the discrete "collocation" inner product.

Now, let $\mathbf{u} = \{u^1, u^2, ..., u^N\}$ nominally represent the values of u at the grid points $\{x_1, x_2, ..., x_N\}$ and expand functions u and v as $u \approx u^l C_l$ and $v \approx v^j C_j$. Then

$$\begin{aligned}(\mathscr{I}(u)v)|_{x=x_i} &\approx (\mathscr{I}(u^l C_l)v^j C_j)|_{x=x_i} \\ &= v^j(\mathscr{I}(u^l C_l)C_j(x))|_{x=x_i} = v^j J_j^i(\mathbf{u}),\end{aligned} \tag{41}$$

where $J_j^i(\mathbf{u}) = \mathscr{I}(u^l C_l)C_j(x)|_{x=x_i}$ is the standard collocation approximation to the linear operator $\mathscr{I}(u)$ based on the cardinal functions $C_j(x)$. The collocation bracket (subscript "c") is then defined as

$$[\mathscr{F}, \mathscr{G}]_c = \frac{\partial \mathscr{F}}{\partial u^i} J^{ij}(\mathbf{u}) \frac{\partial \mathscr{G}}{\partial u^j}, \tag{42}$$

where

$$J^{ij}(\mathbf{u}) = J_k^i w^{jk}. \tag{43}$$

On the other hand, if the weight functions $C^i(x)$ are put into the continuum bracket and the quadrature formula is used to estimate the integral there results

$$\begin{aligned}\{\{C^i, C^j\}\} &\approx \sum_k w_k C^i(x_k).\mathscr{I}(u^l C_l)C^j(x)|_{x=x_k} \\ &= \sum_k w_k C^i(x_k) J_j^k(\mathbf{u})(w^j)^{-1} \\ &= \sum_k w_k (w^i)^{-1}\, \delta_{ik} J_j^k(\mathbf{u})(w^j)^{-1} \\ &= (w^j)^{-1} J_j^i(\mathbf{u}) = J^{ij}(\mathbf{u}) = [u^i, u^j]_c.\end{aligned} \tag{44}$$

So the structure functions $J^{ij}(\mathbf{u})$ of the collocation bracket, Eq. (42), are simply approximations to the exact brackets between the weight functions C^j. This is consistent with Eq. (28) for the weighted residuals method; the structure functions of the bracket are equivalent to simply computing the exact bracket between the approximate delta functions C^i, but using the quadrature rule to approximate the inner product.

From the bracket and an approximation $\mathscr{H}$ to the Hamiltonian we can then write down a set of discrete evolution equations

$$\frac{du^i}{dt} = J^{ij}\frac{\partial \tilde{\mathscr{H}}}{\partial u^j}. \tag{45}$$

However, in contrast to the weighted residuals methods of Section 3, the anti-symmetry of the collocation bracket is not automatic; the use of the quadrature rule and collocation approximation to $\mathscr{I}$ could destroy the anti-symmetry. This must be checked for each method under consideration, if conservation of the Hamiltonian is to be preserved.

CP 4.1. *In order for the collocation bracket to be anti-symmetric it must be that for each i and j*

$$w_i J_j^i(\mathbf{u}) = -w_j J_i^j(\mathbf{u}) \quad \text{(no sum)}.$$

In this case the Hamiltonian will be conserved.

Proof. If $J^{ij}(\mathbf{u}) = -J^{ji}(\mathbf{u})$ then the Hamiltonian will be conserved since the collocation bracket will then be anti-symmetric. But $J^{ij}(\mathbf{u}) = J_k^i(\mathbf{u})w^{jk} = (w_j)^{-1}J_j^i(\mathbf{u})$ (no sum), so for anti-symmetry we require $(w_j)^{-1}J_j^i(\mathbf{u}) = -(w_i)^{-1}J_i^j(\mathbf{u})$, or $w_i J_j^i(\mathbf{u}) = -w_j J_i^j(\mathbf{u})$ for each i and j. ∎

One remark worth making here is that in collocation methods certain degrees of freedom u^l can be associated with physical system boundaries at which, depending on the problem being modeled, an exchange of conserved quantities could take place with some external environment. A system will have the Hamiltonian as a constant of the motion only when it is isolated; when it is not we should expect the Hamiltonian (and other "constants") to change because of the exchange across these boundaries with the rest of the universe. I do not expect that the anti-symmetry condition $w_i J_j^i(\mathbf{u}) = -w_j J_i^j(\mathbf{u})$ should hold at points x_i on the boundary. Indeed, it is possible that Eq. (45) will not apply at such points (e.g., inflow boundaries where u is determined by boundary conditions and not the evolution equation). We shall see some examples of this later.

While the collocation bracket is somewhat less forgiving than the weighted residuals bracket, requiring an explicit check on its anti-symmetry, it still has two nice conservation properties for Casimirs (even, in fact, when the bracket is not anti-symmetric).

CP 4.2. *Suppose that $\mathscr{C}(u) = \langle c, u\rangle$ is a linear Casimir, and suppose that $c(x)$ is interpolated exactly, $c(x) = \sum_i c(x_i)C_i(x)$. Then $\tilde{\mathscr{C}} = \sum_k w_k c(x_k)u^k$ is a Casimir of the collocation bracket.*

Proof. Since $\mathscr{C}$ is a Casimir it must be that $\mathscr{I}(u^l C_l)c = 0$. Therefore, $0 = \mathscr{I}(u^l C_l)c = \mathscr{I}(u^l C_l)\sum_j c(x_j)C_j(x) = \sum_j c(x_j)\mathscr{I}(u^l C_l)C_j(x)$, and so $0 = \sum_j c(x_j)J_j^i(\mathbf{u})$ for all mesh points i. This then implies that

$$[\mathscr{G}, \tilde{\mathscr{C}}]_c = \frac{\partial \mathscr{G}}{\partial u^i}J^{ij}(\mathbf{u})\frac{\partial \tilde{\mathscr{C}}}{\partial u^j} = \frac{\partial \mathscr{G}}{\partial u^i}\sum_j J^{ij}(\mathbf{u})w_j c(x_j)$$

$$= \frac{\partial \mathscr{G}}{\partial u^i}\sum_j J_j^i(\mathbf{u})c(x_j) = 0,$$

which shows that $\tilde{\mathscr{C}}$ is a Casimir of the collocation bracket. ∎

CP 4.3. *Suppose that there is a Casimir $\mathscr{C}(u) = (1/2)\langle u, u\rangle$. Then $\tilde{\mathscr{C}} = \sum_k w_k (u^k)^2$ is a Casimir of the collocation bracket.*

Proof. In this case we have $\delta\mathscr{C}/\delta u = u$ and so it must be that $\mathscr{I}(u^l C_l)u^j C_j(x) = 0$. But this implies that $0 = u^j(\mathscr{I}(u^l C_l)C_j(x))|_{x=x_i} = u^j J_j^i(\mathbf{u})$. Therefore,

$$[\mathscr{G}, \tilde{\mathscr{C}}]_c = \frac{\partial \mathscr{G}}{\partial u^i}\sum_j J^{ij}(\mathbf{u})w_j u^j = \frac{\partial \mathscr{G}}{\partial u^i}\sum_j J_j^i(\mathbf{u})u^j = 0$$

and so $\tilde{\mathscr{C}}$ is a Casimir of the collocation bracket. ∎

Useful results for momentum conservation are not easily derived because they depend on how the Hamiltonian is approximated (which might be accomplished in several ways). The obvious choice of $\mathscr{H}(u^l C_l)$ as the Hamiltonian—simply substituting the finite expansion into the exact Hamiltonian—for use with a collocation bracket could result in tightly coupled convolution sums due to nonlinear terms, just as it can in the weighted residuals method. In order to gain any benefit from the collocation bracket over the weighted residuals method of Section 3 it is necessary to pay some attention to the treatment of nonlinear (really cubic and higher) terms in the Hamiltonian. An example will be presented below.

4.1. *Example: Fourier Collocation*

In order to illustrate this let us develop a Fourier spectral collocation method for the KdV equation with periodic boundary conditions. The exact bracket is again defined by $\mathscr{I} = \partial/\partial x$, so let

$$J_j^i = \begin{cases} (1/2)(-1)^{i+j}\cot((i-j)\pi/N), & i \neq j, \\ 0, & i = j \end{cases} \tag{46}$$

be the standard spectral collocation first differentiation matrix [14] for the points $x_i = i(2\pi/N)$, N even, $i = 0, 1, 2, \ldots, N-1$, and let $w_i = 2\pi/N$ be the trapezoidal rule weights. The matrix J_j^i is clearly anti-symmetric, and the weights are independent of the index, so CP 4.1 is satisfied and the collocation bracket will be anti-symmetric. The resulting bracket will also have the functional $\mathscr{N} = (2\pi/N)\sum_i u^i$ as a Casimir because the Fourier collocation exactly interpolates the constant function (CP 4.2). Ap-

proximating the integral (using the quadrature rule) and derivatives (using the collocation derivative) in the KdV Hamiltonian, Eq. (18), results in a natural approximate Hamiltonian

$$\tilde{\mathscr{H}}_{\mathrm{KdV}} = \frac{2\pi}{N}\sum_{k=0}^{N-1}\left(\frac{1}{6}(u^k)^3 - \frac{1}{2}\left(\sum_{l=0}^{N-1} D_x(k,l)u^l\right)^2\right), \quad (47)$$

where $D_x(k, l) = J_l^k$ is written to emphasize that this is a collocation derivative in x. Although other approximate Hamiltonians could be used, this approximation results in a derivative of $\tilde{\mathscr{H}}$ with respect to the degrees of freedom,

$$\frac{d\tilde{\mathscr{H}}_{\mathrm{KdV}}}{du^j} = \frac{\pi}{N}(u^j)^2 + \frac{2\pi}{N}\sum_{k=0}^{N-1} D_x(k,j)\sum_{l=0}^{N-1} D_x(k,l)u^l, \quad (48)$$

which contains no convolution sums of nonlinear terms. From the approximate Hamiltonian and Eq. (42) we then have a collocation method,

$$\frac{du^i}{dt} = \sum_{j=0}^{N-1} D_x(i,j)\left[\frac{1}{2}(u^j)^2 - \sum_{k=0}^{N-1} D_x(k,j)\sum_{l=0}^{N-1} D_x(k,l)u^l\right] \quad (49)$$

$$= \sum_{j=0}^{N-1}\frac{1}{2}D_x(i,j)(u^j)^2 + \sum_{l=0}^{N-1} D_{xxx}(i,l)u^l, \quad (50)$$

for KdV, in which $D_x(i, j) = J_j^i$ is the usual first-order collocation derivative matrix and $D_{xxx} = D_x^3$ is the usual collocation third derivative matrix. Note that the antisymmetry of $J_j^i = D_x(i, j)$ has been used in going from the first to the second form of this equation. The second form of this method is one that we might write down straight away; it is a totally standard spectral collocation with the advective term treated in a conservative fashion [14, 15]. But this latter form obscures the bracket structure which lies behind the equation and thus it hides the conservation of the Hamiltonian. For Eqs. (49) and (50) do exactly conserve the Hamiltonian $\tilde{\mathscr{H}}_{\mathrm{KdV}}$ and also exactly conserve the Casimir $\tilde{\mathscr{N}}$. This can be confirmed by direct calculation, but by derivation of the method as a series of approximations to the true bracket and Hamiltonian, with the necessary properties of the bracket maintained at each step, these conservation laws become automatic.

This approach, based on discretization of the inner product and $\mathscr{I}$ operator and approximation of the exact Hamiltonian by a quadrature formula, actually underlies the method used by de Frutos *et al.* [16] for the good Boussinesq system, but in that paper the authors simply wrote down the numerical bracket and approximate Hamiltonian without suggesting that the idea in fact has applicability beyond their specific problem. Their goal, successfully achieved, was a fully discrete symplectic method, rather than a conservative method, and in this they benefited from the trivial form of the exact bracket (in particular its independence of u). When a conservative method is sought this is by no means a necessary assumption, as we shall see in Section 5. Again, however, as in Section 3.1, I think that the important point is that a standard numerical method for a Hamiltonian problem can have a Hamiltonian structure of its own, described by a numerical bracket and an approximate Hamiltonian.

4.2. *Legendre–Gauss–Lobatto Collocation*

As another example of the general method of derivation, let us examine the nonlinear advection equation, but without the assumption of periodic boundary conditions, by using Legendre–Gauss–Lobatto collocation on the interval $-1 \le x \le 1$. In this case the collocation points x_i, $i = 0, 1, \ldots, N$, will be the $N + 1$ zeros of the polynomial $(1 - x^2)\, dP_N/dx$, where P_N is the Nth Legendre polynomial, and the quadrature weights will be $w_i = 2/(N(N + 1) P_N^2(x_i))$ [17]. The cardinal functions are

$$C_i(x) = \frac{(1-x^2)\, dP_N/dx}{N(N+1)P_N(x_i)}\frac{1}{x_i - x} \quad (51)$$

(as is trivially verified from the differential equation $d((1 - x^2)\, dP_N/dx)/dx + N(N + 1)P_N = 0$) and the collocation derivative is then given by the matrix

$$J_j^i = D_x(i,j) = \begin{cases} \dfrac{P_N(x_i)}{P_N(x_j)}\dfrac{1}{x_i - x_j}, & i \ne j, \\ -N(N+1)/4, & i = j = 0, \\ N(N+1)/4, & i = j = N, \\ 0, & i = j,\ i, j \ne N, 0. \end{cases} \quad (52)$$

The J_j^i and weights w_i obviously satisfy $w_i J_j^i = -w_j J_i^i$, for all i, j except $i = j = 0$ or $i = j = N$ (that is, everywhere except the boundaries); this is an inevitable consequence of the fact that Legendre–Gauss–Lobatto quadrature can exactly integrate a polynomial of order $2N - 1$. The failure of the anti-symmetry condition in the bracket at the boundaries is physically correct because the system is not isolated; there are boundary terms that arise from integration by parts in the true bracket

$$\{\{\mathscr{F}, \mathscr{G}\}\} = \int_{-1}^{1}\frac{\delta\mathscr{F}}{\delta u}\frac{\partial}{\partial x}\frac{\delta\mathscr{G}}{\delta u}\, dx$$
$$= -\int_{-1}^{1}\frac{\delta\mathscr{G}}{\delta u}\frac{\partial}{\partial x}\frac{\delta\mathscr{F}}{\delta u}\, dx + \left[\frac{\delta\mathscr{F}}{\delta u}\frac{\delta\mathscr{G}}{\delta u}\right]_{-1}^{1}. \quad (53)$$

Thus, the true bracket is not really anti-symmetric because of these boundary terms, and the same is true of the collocation bracket. So in this case CP 4.1 will apply and the Hamiltonian will be conserved, except for boundary terms. Similarly, because Nth-order polynomial interpolation can exactly expand a constant, CP 4.2 applies, and a method derived from the Legendre–Gauss–Lobatto collocation bracket will have the functional $\mathscr{N} = \sum_i w_i u^i$ as a Casimir; we can therefore conclude that this quantity will be a constant in the discrete method, except for boundary terms.

Let us briefly examine these properties of a Legendre–Gauss–Lobatto collocation method for the nonlinear advection equation. The Hamiltonian can be approximated by $\tilde{\mathscr{H}}_{\mathrm{NA}} = (1/6)\sum_{i=0}^{N} w_i (u^i)^3$. The structure functions for the collocation bracket are $J^{ij} = (w_j)^{-1} J^i_j$, and so the semi-discrete equation becomes

$$\frac{du^i}{dt} = J^{ij}\frac{\partial \tilde{\mathscr{H}}_{\mathrm{NA}}}{\partial u^j} = \frac{1}{2}\sum_{j=0}^{N} D_x(i,j)(u^j)^2, \tag{54}$$

which is nothing unusual—it is a rather standard collocation method for the nonlinear advection equation written in conservative form. As such, it is not surprising that the Casimir $\tilde{\mathscr{N}} = \sum_{i=0}^{N} w_i u^i$ is conserved, in the sense that

$$\frac{d\tilde{\mathscr{N}}}{dt} = \frac{1}{2}\sum_{j=0}^{N}\sum_{i=0}^{N} w_i J^i_j (u^j)^2 = \frac{1}{2}(u^N)^2 - \frac{1}{2}(u^0)^2, \tag{55}$$

which is as it should be. It is less obvious from Eq. (54) and $\mathscr{H}_{\mathrm{NA}}$ is a constant, but this is obvious from the anti-symmetric structure of the collocation bracket. A direct calculation is therefore unnecessary, but easy because of the anti-symmetry condition $w_i J^i_j = -w_j J^j_i$ for all i and j not on the boundary, and yields

$$\frac{d\tilde{\mathscr{N}}_{NA}}{dt} = \frac{1}{4}\sum_{i=0}^{N}\sum_{j=0}^{N} (u^i)^2 w_i J^i_j (u^j)^2 = \frac{1}{8}(u^N)^4 - \frac{1}{8}(u^0)^4. \tag{56}$$

Thus, two of the conservation properties of the semi-discrete method, Eq. (54), are natural consequences of the numerical bracket which has been inherited by the collocation method from the Poisson bracket for the exact system.

5. COLLOCATION METHODS FOR THE VLASOV–MAXWELL EQUATIONS

The previous examples have all been based on the trivial bracket defined by $\mathscr{J} = \partial/\partial x$. A more complex bracket structure underlies the Vlasov–Maxwell [18] equations of plasma kinnetic theory, and in this section I shall use this to further illustrate the ideas about numerical brackets described above. In one spatial dimension the Vlasov–Maxwell equations are

$$\frac{\partial f_\alpha}{\partial t}(x,v,t) = -v\frac{\partial f_\alpha}{\partial x}(x,v,t) - \frac{q_\alpha}{m_\alpha}E(x,t)\frac{\partial f_\alpha}{\partial v}(x,v,t) \tag{57}$$

$$\frac{\partial E}{\partial t}(x,t) = -4\pi\sum_\alpha q_\alpha \int_{-\infty}^{\infty} v f_\alpha(x,v,t)\,dv, \tag{58}$$

where v represents an independent variable—physically the velocity of particles—and α is a species index denoting a charged particle species whose mass is m_α and whose charge is q_α. These equations relate the density f_α of charged particle species in the single particle phase space (x, v) to their self-consistent electric field, E. Physically the distribution functions f_α and electric field E are constrained by Coulomb's law

$$\frac{\partial E}{\partial x}(x,t) = 4\pi\sum_\alpha q_\alpha \int_{-\infty}^{\infty} f_\alpha(x,v,t)\,dv. \tag{59}$$

The physical boundary conditions that often apply to f_α are $f_\alpha \to 0$ rapidly as $|v| \to \infty$, $f_\alpha(-1, v, t)$ given for $v > 0$, and $f_\alpha(1, v, t)$ given for $v < 0$; these boundary conditions in x say that the distribution of particles flowing into the region between $x = -1$ and $x = 1$ is specified. The boundaries could be spatial locations other than $x = \pm 1$, of course; this is simply a trivial matter of scaling and shifting. In the development that follows I shall truncate the velocity range and scale it into $-1 \le v \le 1$ as well; scale factors that result from such a scaling will not be written, however, as they only serve to dirty the page. Physically a boundary condition at $v = \pm 1$ should specify the distribution functions f_α at "inflow" points, determined by the sign of $q_\alpha E(x, t)$ at each point. In simulations one would most likely choose the maximum and minimum velocity for truncation based on a knowledge of the fastest possible particle in the system being simulated so that $f_\alpha(x, \pm 1, t) = 0$. If this is true initially, at $t = 0$, then it will be true for at least some finite time range $0 \le t \le T$ because particles are subject to finite accelerations.

The bracket for these equations, presented in a trivially different form in Ref. [1], is

$$\begin{aligned}\{\{F, G\}\} = &\sum_\alpha \int_{-1}^{1}\int_{-1}^{1} \frac{1}{m_\alpha}\frac{\delta F}{\delta f_\alpha}(x,v)\left\{\frac{\delta G}{\delta f_\alpha}, f_\alpha\right\}(x,v)\,dx\,dv \\ &+ 4\pi\sum_\alpha \frac{q_\alpha}{m_\alpha}\int_{-1}^{1}\int_{-1}^{1}\frac{\delta F}{\delta E}(x)\frac{\partial f_\alpha}{\partial v}(x,v)\frac{\delta G}{\delta f_\alpha}(x,v) \\ &- \frac{\delta G}{\delta E}(x)\frac{\partial f_\alpha}{\partial v}(x,v)\frac{\delta F}{\delta f_\alpha}(x,v)\,dx\,dv,\end{aligned} \tag{60}$$

where

$$\left\{\frac{\delta G}{\delta f}, f\right\} = \frac{\partial f}{\partial v}\frac{\partial}{\partial x}\frac{\delta G}{\delta f} - \frac{\partial f}{\partial x}\frac{\partial}{\partial v}\frac{\delta G}{\delta f} \tag{61}$$

is the canonical single particle Poisson bracket. The Hamiltonian for the Vlasov–Maxwell dynamics is the sum of the particle and field energies:

$$\mathscr{H}(f, E) = \sum_\alpha \int_{-1}^{1}\int_{-1}^{1} \frac{m_\alpha v^2}{2} f_\alpha \, dx \, dv + \int_{-1}^{1} \frac{1}{8\pi} E^2 \, dx. \tag{62}$$

Some of the more interesting constants of the motion for the Vlasov–Maxwell system besides the Hamiltonian include the number of particles in a species,

$$\mathscr{N}_\alpha = \int_{-1}^{1}\int_{-1}^{1} f_\alpha \, dx \, dv, \tag{63}$$

which is a Casimir, and total linear momentum of the particles,

$$\mathscr{M} = \sum_\alpha \int_{-1}^{1}\int_{-1}^{1} m_\alpha v f_\alpha \, dx \, dv, \tag{64}$$

which is a momentum in the formal sense, corresponding to the invariance of the Hamiltonian under spatial translations. It is important to note, however, that the momentum is a constant only if the constraint given by Eq. (59) holds; formally, $\mathscr{M}$ only generates the group of translations (acting by $f(x, v) \mapsto f(x - t, v)$, $E(x) \mapsto E(x - t)$) when Eq. (59) is satisfied. Equation (59), Coulomb's Law, implies that the force between two particles depends only on the distance between them, and their location with respect to some external reference frame is immaterial. This means that the center of mass of the system will feel no force and will simply translate at a constant velocity. Another important set of constants is the Casimirs

$$\mathscr{S}_\alpha = \int_{-1}^{1}\int_{-1}^{1} f_\alpha^2 \, dx \, dv, \tag{65}$$

whose preservation is attractive in part because it provides stability for the numerical method (which the physical energy does not, in the absence of a guarantee of positiveness of the numerical approximations to the distribution functions).

I have previously described a method [19] and presented numerical results based on Legendre–Gauss–Lobatto collocation for the Vlasov–Maxwell equations. That method conserved energy, particles, and momentum even for discrete time, but it was not noted there that the preservation of these conservation laws could be understood in terms of the bracket structure of the system. In this section I wish to make this connection clear.

The collocation points in x will be taken to be the $N_x + 1$ roots $x_0, x_1, ..., x_{N_x}$ of the polynomial $(1 - x^2)\, dP_{N_x}/dx$, and the collocation points in v will be the roots $v_0, v_1, ..., v_{N_v}$ of $(1 - v^2)\, dP_{N_v}/dv$; all the scale factors that arise by mapping the problem domain into $-1 \le x \le 1$ and $-1 \le v \le 1$ can be absorbed into the quadrature weights and collocation derivative matrix elements. The collocation derivative matrices in x and v will be denoted $D_x(i, k)$ and $D_v(j, l)$, respectively; except possibly for trivial scaling factors, these matrices are given by Eq. (52). The quadrature weights for x and v integration will similarly be denoted w_i^x and w_j^v. As mentioned in Section 4.2, the Legendre–Gauss–Lobatto weights and collocation derivative satisfy the anti-symmetry property of CP 4.1, except at the system boundary (where we do not want them to); this will be enough to make the collocation bracket, written out explicitly below, anti-symmetric except for boundary terms. Therefore, whatever approximate Hamiltonian we use to define the dynamics, it will be conserved, except for inflow and outflow at the system boundaries. Furthermore, both CP 4.2 and CP 4.3 will be satisfied for $\mathscr{N}_\alpha$ and S_α, respectively, so a method based on Legendre–Gauss–Lobatto collocation will conserve particles and the square integral of the distribution function, again, except for physically correct inflow and outflow.

Because of the large number of indices now required (for space x, velocity v, and species) I shall dispense with the summation convention; the only sums are those explicitly written, and the placement of indices will be dictated by readability. The structure functions of the collocation bracket between two distribution function values f_α^{ij} and f_β^{kl} is determined by the first double integral in Eq. (60). This term for species α generates a piece of $\mathscr{I}$ that looks like

$$\mathscr{I}(f_\alpha) = \frac{\partial f_\alpha}{\partial u}\frac{\partial}{\partial x} - \frac{\partial f_\alpha}{\partial x}\frac{\partial}{\partial u}. \tag{66}$$

The structure function between f_α^{ij} and f_β^{kl} is zero if $\alpha \ne \beta$; otherwise,

$$[f_\alpha^{ij}, f_\alpha^{kl}]_c = \frac{1}{m_\alpha w_k^x w_l^v}\left[\delta_{lj} D_x(i, k) \sum_{m=0}^{N_v} D_v(j, m) f_\alpha^{im} - \delta_{ki} D_v(j, l) \sum_{n=0}^{N_x} D_x(i, n) f_\alpha^{nj}\right]. \tag{67}$$

These are most easily computed by putting the weight functions into the exact bracket and replacing the sums by the quadrature rule, as described in Eq. (44). Alternately, the collocation approximation to the operator $\mathscr{I}(f_\alpha)$ in Eq. (66) can be used and divided by the quadrature weights w_k^x and w_l^v as in Eq. (43). Note again that these structure functions are anti-symmetric only at interior points; at

boundary points, $i = k = 0$, $i = k = N_x$, $j = l = 0$, or $j = l = N_v$, the structure function is not anti-symmetric. This is as it should be so that inflow and outflow are accounted for. The bracket between the field and the distribution function comes from the second double integral in Eq. (60), which generates more species specific pieces of $\mathscr{I}$ of the form

$$\mathscr{I}(f_\alpha)\begin{pmatrix}\dfrac{\delta\mathscr{G}}{\delta f_\alpha}\\[2ex] \dfrac{\delta\mathscr{G}}{\delta E}\end{pmatrix} = \frac{4\pi q_\alpha}{m_\alpha}\begin{pmatrix} -\dfrac{\partial f_\alpha}{\partial v}\dfrac{\delta\mathscr{G}}{\delta E} \\[2ex] \displaystyle\int_{-1}^{1}\frac{\partial f_\alpha}{\partial v}\frac{\delta\mathscr{G}}{\delta f_\alpha}\,du\end{pmatrix}. \tag{68}$$

Using the collocation approximation to this operator, and again raising the second index, yields the structure functions

$$[f_\alpha^{ij}, E^k]_c = -\delta_{ik}\frac{4\pi}{w_k^x}\frac{q_\alpha}{m_\alpha}\sum_{m=0}^{N_v} D_v(j,m) f_\alpha^{im} \tag{69}$$

$$[E^k, f_\alpha^{ij}]_c = \delta_{ki}\frac{4\pi}{w_i^x}\frac{q_\alpha}{m_\alpha}\sum_{m=0}^{N_v} D_v(j,m) f_\alpha^{km}. \tag{70}$$

These terms are anti-symmetric even at boundary points because this part of the exact bracket is explicitly anti-symmetric, independent of the details at the boundary.

For the Vlasov–Maxwell equations a natural approximation to the Hamiltonian is

$$\tilde{\mathscr{H}} = \sum_\alpha \sum_{\substack{i=0\\ i=N_x}}^{\substack{j=0\\ j=N_v}} \frac{m_\alpha v_j^2}{2} f_\alpha^{ij} w_i^x w_j^v + \frac{1}{8\pi}\sum_{i=0}^{N_x}(E^i)^2 w_i^x. \tag{71}$$

From this we can compute the derivatives

$$\frac{\partial\tilde{\mathscr{H}}}{\partial f_\alpha^{ij}} = \frac{m_\alpha v_j^2}{2} w_i^x w_j^v \tag{72}$$

and

$$\frac{\partial\tilde{\mathscr{H}}}{\partial E^i} = \frac{1}{4\pi} E^i w_i^x. \tag{73}$$

The quadrature weights in these derivatives will be canceled out by their inverses appearing in the structure functions of the numerical collocation bracket, and the discrete version of the Vlasov–Maxwell equations will then be

$$\frac{df_\alpha^{ij}}{dt} = -v_j\sum_{k=0}^{N_x} D_x(i,k) f_\alpha^{kj} - \frac{q_\alpha}{m_\alpha} E^i \sum_{l=0}^{N_v} D_v(j,l) f_\alpha^{il} \tag{74}$$

$$\frac{dE^i}{dt} = -4\pi\sum_\alpha q_\alpha \sum_{j=0}^{N_v} w_j^v v_j f_\alpha^{ij} \tag{75}$$

which is the form that was written down directly as a collocation approximation in Ref. [19]. In writing these, it has been noted that the collocation derivative exactly differentiates a constant in x, $\sum_n D_x(i,n) = 0$, and a quadratic in v, $\sum_m D_v(j,m) v_m^2/2 = v_j$, assuming only that $N_x \geq 1$ and $N_u \geq 2$. Equations (74) and (75) correctly conserve $\tilde{H}$, $\tilde{\mathscr{N}}_\alpha = \sum_{ij} w_i^x w_j^v f_\alpha^{ij}$, and $\tilde{\mathscr{J}}_\alpha = \sum_{ij} w_i^x w_j^v (f_\alpha^{ij})^2$ because of CP 4.1, CP 4.2, and CP 4.3, again provided that $N_x \geq 1$ and $N_u \geq 2$, and of course, boundary terms will appear to account for inflow and outflow. These conservation properties were observed in the simulation described in Ref. [19], which made use of a second-order Runge–Kutta time discretization; understanding systematically what properties of this discretization were responsible for these conservation properties was in fact part of the motivation for the work described in this paper.

Momentum conservation is also preserved. Note that because the equations conserve particles they are still consistent with Coulomb's law; that is, if the discrete Coulomb's law

$$\sum_{k=0}^{N_x} D_x(i,k) E^k = 4\pi\sum_\alpha q_\alpha \sum_{j=0}^{N_v} w_j^v f_\alpha^{ij} \tag{76}$$

is satisfied initially, then it will be satisfied for all time, except for boundary terms arising from particles accelerating out of the system at the velocity boundaries $v = \pm 1$. Using this it can be directly checked that $[\tilde{\mathscr{M}}, \tilde{\mathscr{H}}]_c$ is given purely by boundary terms, indicating that momentum is constant, except for the physical loss of momentum through the system boundaries.

The nonconservation of energy in a plasma physics simulation is usually called grid heating or self-heating, while the nonconservation of momentum implies the existence of a net grid force. The method described here, based on the Hamiltonian structure of the equations, has neither grid heating nor a net grid force. The only reason energy, momentum, particles, or $\mathscr{S}_\alpha$ is not exactly constant is that particles can enter and leave the system through its physical boundaries.

There have been other methods for plasma kinetic simulation which conserve energy, and Lewis [20, 21] has in fact described a class of Lagrangian (characteristic or particle-following) methods that are derived from the principle of least action. Through the introduction of suitable generalized momenta this naturally results in a system of equations that are in canonical Hamiltonian form. More recently Scovel and Weinstein [22] have developed a class of methods that are based on a numerical bracket that satisfies the Jacobi identity. Both of these methods are essentially

particle methods and, as such, rather different in character from the Eulerian discretizations described here. Scovel and Weinstein [22] in fact suggest that it is not possible to develop a discretization of the Vlasov–Poisson system that inherits a good (Jacobi) Poisson bracket by using discretizations on grids in x and v, but that instead, particle methods must be used. Nothing in the present paper suggests otherwise, since the bracket of Eqs. (67), (69), and (70) does not satisfy the Jacobi condition. The relative advantages of particle methods of plasma simulation versus the so-called ''Vlasov solvers'' are probably problem dependent; particle methods have great advantages in cold plasmas, but can yield noisy distributions and electric fields, especially in those regions of phase space that contain relatively few particles. In contrast, Vlasov solvers can provide noise-free resolution of high-velocity tails and other low-density regions, but they can have difficulties resolving the ever-changing scales of an evolving distribution function, especially at long times (at least in the absence of adaptive grids or filtering).

6. CONCLUSIONS

The point is that Hamiltonian partial differential equations give special structure; you should consider this when deriving numerical methods. The goal is not to derive strange, new, exotic numerical methods. Rather it is that very ordinary looking numerical methods can have some of the Hamiltonian structure of the problem trapped inside them. The exact Poisson bracket can be replaced by a numerical bracket and structure functions can be computed from this to provide a discretization of the system, with evolution defined by the numerical bracket and an approximate Hamiltonian. Recognition of this can allow the identification of which constants of the motion have been preserved (in the limit of continuous time) by a particular method.

It is hoped that the explicit construction techniques given here for weighted residuals methods, including collocation methods, will be of use to those interested in examining and understanding the conservation properties of existing methods, or in designing methods that possess desired conservation properties.

REFERENCES

1. P. J. Morrison, *Phys. Lett. A 80*, 383 (1980).
2. J. E. Marsden and A. Weinstein, *Physica D* **4,** 394 (1982).
3. D. D. Holm, J. E. Marsden, T. Ratiu, and A. Weinstein, *Phys. Rep.* **123,** 1 (1985).
4. T. Shepherd, *Adv. Geophys.* **32,** 287 (1990).
5. D. Greenspan, *J. Comput. Phys.* **56,** 28 (1984).
6. J. M. Snaz-Serna, *BIT* **28,** 877 (1988).
7. F. M. Lasagni, *J. Appl. Math. Phys.* **39,** 952 (1988).
8. P. J. Channell and C. Scovel, *Nonlinearity* **3,** 231 (1990).
9. H. Goldstein, *Classical Mechanics* (Addison–Wesley, Reading, MA, 1980).
10. P. Olver, *Applications of Lie Groups to Differential Equations* (Springer-Verlag, Berlin, 1986).
11. J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems* (Springer-Verlag, Berlin, 1994).
12. C. S. Gardner, *J. Math. Phys.* **12,** 1548 (1971).
13. D. Gottlieb and S. A. Orszag, *Numerical Analysis of Spectral Methods: Theory and Applications* (SIAM, Philadelphia, 1977).
14. D. Gottlieb, M. Y. Hussaini, and S. A. Orszag, ''Theory and Applications of Spectral Methods,'' in *Spectral Methods for Partial Differential Equations*, edited by R. G. Voigt, D. Gottlieb, and M. Y. Hussaini (SIAM, Philadelphia, 1984).
15. S. B. Wineberg, J. F. McGrath, E. F. Gabl, L. R. Scott, and C. E. Southwell, *J. Comput. Phys.* **97,** 311 (1991).
16. J. de Frutos, T. Ortega, and J. M. Sanz-Serna, *Comput. Methods Appl. Mech. Eng.* **80,** 417 (1990).
17. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).
18. A. Vlasov, *J. Phys.* **9,** 25 (1945).
19. J. P. Holloway, *Trans. Am. Nucl. Soc.* **69,** 232 (1993).
20. H. R. Lewis, *J. Comput. Phys.* **6,** 136 (1970).
21. H. R. Lewis, *J. Comput. Phys.* **10,** 400 (1972).
22. C. Scovel and A. Weinstein, *Commun. Pure Appl. Math.* **47,** 683 (1994).