Symplectic Integration of Hamiltonian Wave Equations

Robert McLachlan

April 12, 1993

Program in Applied Mathematics, Campus Box 526,
University of Colorado at Boulder, Boulder, Colorado 80309–0526.

Phone: (303)492–5199   Fax: (303)492–4066

`rxm@boulder.colorado.edu`

Abstract

The numerical integration of a wide class of Hamiltonian partial differential equations by standard symplectic schemes is discussed, with a consistent, Hamiltonian approach. We discretize the Hamiltonian and the Poisson structure separately, then form the the resulting ODE's. The stability, accuracy, and dispersion of different explicit splitting methods are analyzed, and we give the circumstances under which the best results can be obtained; in particular, when the Hamiltonian can be split into linear and nonlinear terms. Many different treatments and examples are compared.

Mathematical Subject Classification: 65M20, 58F05, 35L70, 35Q20

0. Introduction

Symplectic integrators have a remarkable capacity for capturing the long-time dynamics of Hamiltonian systems correctly and easily. Invariant sets and orbit statistics converge even if the orbit is not followed with pointwise accuracy (MacKay (1992), McLachlan and Atela (1992), Candy and Rozmus (1991)). Indeed, if one wants to compute $u(T)$ accurately for fixed T then conventional methods may do better because they can have smaller local truncation errors. This is not an appropriate test for symplectic integrators: one should concentrate instead on phase space structures (e.g. the shape of a traveling wave) and not on temporal errors (its speed). Whether or not accumulating phase errors (e.g. of angles on invariant tori) corrupt the dynamics depends on the particular system studied, and on the measured property being structurally stable in the space of Hamiltonian systems.

The potential of symplectic integrators is even greater for *partial* differential equations. Now the equations are stiff—they contain widely different time-scales. Hence (when the Courant number is order 1) a non-conservative scheme will introduce a lot of dissipation in the high modes; but standard conservative schemes (Crank-Nicolson, 3-time-level leapfrog) can have undesirable attributes such as implicitness, parasitic waves, or low order. Hamiltonian methods have no parasitic waves, extend easily to any order, and are often explicit.

The recent literature has many examples of *ad hoc* methods developed for particular Hamiltonian PDE's, with an emphasis on proofs of convergence of $u(x,T)$ for fixed T and of stability with respect to initial data. Whether such proofs can be carried through depends on the precise form of the nonlinear terms. There are energy-conserving schemes (Fei and Vásquez (1991), the sine-Gordon equation; Glassey (1992), the Zakharov equations; Glassey and Schaeffer (1991), a nonlinear wave equation) and unconditionally stable schemes (Dai (1992), the variable-coefficient Schrödinger equation; de Frutos, Ortega and Sanz-Serna (1991), the Boussinesq equation; Ivanauskas (1991), nonlinear Schrödinger equations). Only this last is symplectic—it is just the midpoint rule (see section 2). De Frutos and Sanz-Serna (1991) discuss the (implicit) midpoint rule and extensions.

Here we take a more general approach and study the characteristics of a wide class of methods applied to general Hamiltonian PDE's. Instead of discretizing the PDE directly, we discretize both the Hamiltonian function and the Hamiltonian (Poisson) structure, then form the resulting ODE's. This is the approach taken, for example, by Feng and Qin (1987) for the linear wave equation, and by de Frutos and Sanz-Serna (1990) for the Boussinesq equation. As long as the Hamiltonian structure is canonical (see e.g. eq. (1.1)) and the symmetries of the derivatives are preserved in the difference operators, then standard methods will arise in this way (as in §5); so this approach is only strictly necessary in the non-canonical case (see examples 2 and 3, and §7).

Section 1 outlines the structure of Hamiltonian PDE's and section 2 reviews symplectic integration as needed here. We concentrate on the the splitting methods of Ruth (1983), Suzuki (1991), and Yoshida (1990). These may be applied when the Hamiltonian is separable ($H = T(p)+V(q)$, "P-Q" splitting), but the best methods are possible when the Hamiltonian is a sum of linear and nonlinear terms, each of which can be integrated exactly ("L-N" splitting). Wave equations (Strauss, 1989) (linear equations with periodic solutions, to which a nonlinear term is added) are often of this form. Proposition 1 proves sufficient conditions for the more accurate Runge-Kutta-Nyström methods to apply without modification to this linear+nonlinear splitting. Local truncation errors are easily computed using Poisson brackets of the parts of the Hamiltonian.

The behavior of the high-frequency modes is central, and affects the stability of a method and dominates dispersion errors. In contrast to stiff dissipative systems (the heat equation) these modes must almost always be temporally resolved for successful calculations. Schemes which try to avoid doing so either lose accuracy or are not efficient. Section 3 compares the stability, accuracy and dispersion of the different symplectic integrators, particularly with regard to P-Q splitting. Section 4 deals with L-N splitting: If the equation is linear in its highest derivatives, dispersion errors are almost eliminated. We prove that (under the right conditions) two time-steps per period of the fastest wave are sufficient for stability; in many examples they turn out to be necessary as well.

We give examples throughout and compare the Hamiltonian approach to other published schemes. Examples (all wave equations) are the Boussinesq, Korteweg-de Vries, Zakharov, and sine-Gordon equations; three more are treated in greater depth in sections 5–7. Section 5 illustrates spectral discretization on a nonintegrable nonlinear wave equation and compares different symplectic schemes. We show how aliasing errors may be eliminated to any desired degree within the Hamiltonian framework, and how this can help in marginally-resolved calculations. We also compare symplectic schemes to an energy-preserving one—it turns out that *reversibility* is more important than exact energy conservation.

Section 6 deals with the (integrable) nonlinear Schrödinger equation. Ablowitz and Herbst (1990) have studied this equation and concluded that an integrable spatial discretization is essential in following orbits close to homoclinic. Instead, we find that sufficient spatial resolution and symplectic time-stepping are the essential ingredients for a successful long-time integration; accuracy and integrability are not crucial.

Section 7 compares three different fourth-order symplectic integrators applied to the Korteweg-de Vries equation, with regard to accuracy and speed.

1. Hamiltonian partial differential equations

Olver (1986) is a good introduction to the structure of Hamiltonian ODE's and PDE's. Here we give a brief overview. A Hamiltonian dynamical system consists of a triple $(M, \{\cdot,\cdot\},$

H) where M is a smooth manifold (the *phase space*), $H : M \to \mathbb{R}$ is the Hamiltonian function, and $\{\cdot,\cdot\}$ is a Poisson bracket, a bilinear, skew-adjoint operator $C(M)\times C(M) \to \mathbb{R}$ satisfying the Jacobi identity

$$\{F,\{G,H\}\} + \{G,\{H,F\}\} + \{H,\{F,G\}\} = 0 \quad \forall F,G,H : M \to \mathbb{R}$$

and the Leibniz rule

$$\{F,GH\} = \{F,G\}H + \{F,H\}G \qquad \forall F,G,H : M \to \mathbb{R}.$$

The bracket can be written in coordinates x_i as

$$\{F,G\} = (\nabla F)^T J(x)(\nabla G)$$

where J is called the Poisson tensor. A change of variables $x \to X = \phi(x)$ induces a bracket in the new variables by

$$\{F\circ\phi, G\circ\phi\}_X = \{F,G\}_x \circ \phi$$

or, in coordinates,

$$(D_x\phi)J(x)(D_x\phi)^T = \widetilde{J}(X).$$

If, in addition, $X \in M$ (i.e. ϕ is a map of phase space into itself, such as the time-step in a numerical integrator) then it is natural to require $\widetilde{J}(x) = J(x)$; in this case ϕ is called a Poisson map. The time-map of the Hamiltonian dynamical system

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

is a Poisson map. A symplectic (or Poisson) integrator is one for which a time step is a Poisson map. Casimirs are functionals C such that $\{C,F\} = 0\ \forall F$, hence integrals of the motion for any H. (Thus they might be called kinematic, rather than dynamic invariants.) The number of Casimirs is corank(J).

When the phase space is infinite dimensional, we write the triple as $(\mathcal{M},\{\cdot,\cdot\},\ \mathcal{H})$, and the Poisson operator as $\mathcal{J}$. Typically $\mathcal{M}$ consists of sets of smooth functions on a finite dimensional space Z, e.g. with periodic boundary conditions, $\mathcal{M} = C^r(\mathbb{T}^n,\mathbb{R}^m)$. An element in $\mathcal{M}$ is $u(x)$, $x \in Z$. The Hamiltonian $\mathcal{H} : \mathcal{M} \to \mathbb{R}$ is a functional on this space, and the bracket can be written as

$$\{\mathcal{F},\mathcal{G}\}[u] = \int_Z \frac{\delta\mathcal{F}}{\delta u}\mathcal{J}(u)\frac{\delta\mathcal{G}}{\delta u}\,dx$$

where $\delta\mathcal{F}/\delta u$ is the variational, or Gateaux, derivative defined by

$$\left(\frac{d}{d\epsilon}F[u+\epsilon v]\right)_{\epsilon=0} \equiv \int_Z \frac{\delta\mathcal{F}}{\delta u}\,v\,dx.$$

When $\mathcal{J}$ is constant on $\mathcal{M}$, the Jacobi identity is trivially satisfied, and one need only check skew-adjointness. There are three primary examples to which we shall refer. First, the *nonlinear wave equation* (NLW), for which $u = (q(x), p(x))^T$, $\mathcal{J}$ is the canonical Poisson operator $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and $\mathcal{H} = \int (\frac{1}{2}p^2 + \frac{1}{2}(q_x)^2 + V(q))\,dx$ giving rise to dynamics

$$\begin{aligned} \dot{q} = \{q, \mathcal{H}\} = & \quad \frac{\delta \mathcal{H}}{\delta p} = p \\ \dot{p} = \{p, \mathcal{H}\} = & -\frac{\delta \mathcal{H}}{\delta q} = q_{xx} - V'(q) \end{aligned} \tag{1.1}$$

One might take Z = the circle $\mathbb{S}^1$ and $\mathcal{M} = C^r(Z, \mathbb{R}^2)$, bearing in mind that for some $u \in \mathcal{M}$ (depending on $V(q)$) the solution might not exist for all time. Second, the *nonlinear Schrödinger equation* (NLS), $\mathcal{J}$ is $-i$ times the canonical one, ψ and ψ^* are canonically conjugate variables (regarded as independent) and $\mathcal{H}(\psi, \psi^*) = \int (|\psi_x|^2 + |\psi|^4)\,dx$, giving

$$i\dot{\psi} = \frac{\delta \mathcal{H}}{\delta \psi^*} = \psi_{xx} + 2|\psi|^2\psi \tag{1.2}$$

and its complex conjugate. (Note that real canonical variables are $(q, p) = (\sqrt{2}\Re\psi, \sqrt{2}\Im\psi)$.) Third, the *Korteweg-de Vries* (KdV) equation: $\mathcal{J} = \partial_x$ and $\mathcal{H} = \int (-u^3 + \frac{1}{2}(u_x)^2)\,dx$:

$$\dot{u} = \partial_x(-3u^2 - u_{xx}). \tag{1.3}$$

In general, we are not interested here in the integrability or otherwise of the PDE's; given $\mathcal{J}$, we wish to consider the widest possible class of Hamiltonians.

To reduce a Hamiltonian PDE to a set of Hamiltonian ODE's which can be symplectically integrated, our approach is to discretize $\mathcal{J}$ and $\mathcal{H}$ separately and then form the resulting dynamical system. (This is in contrast to reductions of integrable PDE's, which are usually found by the method of inverse scattering, the Hamiltonian structure then being found by inspection.) $\mathcal{H}$ is an integral which can be discretized in any (suitably accurate) way, being careful to maintain the symmetry of any derivatives in $\mathcal{H}$. If $\mathcal{J}$ is constant, it may be discretized by replacing the differential operators by any (matrix) difference operator D, for example, central or pseudo-spectral differences. In what follows D will be any appropriate matrix difference operator. Some points to remember are that $\partial_x\partial_x = \partial_{xx}$ usually breaks down when discretized, $\mathcal{H}$ may be integrated by parts as necessary to get compact differences, and that for equations involving odd derivatives we may get 1-point-more-compact differences using staggered grids.

Example 1:
$\mathcal{H} = -\int (u_x)^2\,dx = \int u u_{xx}\,dx$ may be discretized as $H = \sum_{i,j} u_i D_{ij} u_j$ giving $(\nabla H)_i = \sum_j (D_{ij} + D_{ji}) u_j$. This approximates $\frac{\delta \mathcal{H}}{\delta u} = 2u_{xx}$ only if $D_{ij} = D_{ji}$, i.e. the matrix D must be symmetric. So Chebyshev-spectral or finite differences skewed at boundaries are not suitable. ■

There are some non-constant operators for which the same procedure works, examples being $\mathcal{J} = \begin{bmatrix} 0 & \partial_x u_1 \\ u_1 \partial_x & 0 \end{bmatrix}$ and $\mathcal{J} = u\partial_x u$. (Here the operator ∂_x acts to the right, i.e. $\mathcal{J}v = u(uv)_x$ in the latter case.) In these cases the same change of variables which makes $\mathcal{J}$ constant also works for the discrete versions ($(q,p) = (u_2, \exp \frac{\partial u_1}{\partial x})$ makes $\widetilde{\mathcal{J}}$ canonical in the first case, and $q = \frac{1}{2}u^2$ gives $\widetilde{\mathcal{J}} = \partial_x$ in the second). But for many other operators a straightforward discretization of derivatives does *not* work, i.e. the resulting J does not satisfy the Jacobi identity, so the discrete system is not Hamiltonian. For example, one can show that $\mathcal{J} = u\partial_x + \partial_x u$ (which arises in a different formulation of KdV) has no Hamiltonian finite-difference approximations. One can try and change variables to make J constant ($v = \sqrt{2u} \Rightarrow \widetilde{\mathcal{J}} = \partial_x$ does the trick here if $u > 0$), but whether this can be done in general is unresolved at present (see Olver (1988)), and may be unwieldy even if it is possible. Thus, finite-dimensional truncations of non-constant $\mathcal{J}$'s remains an important problem.

Henceforth we restrict our attention to constant $\mathcal{J}$'s, primarily the canonical one and $\mathcal{J} = \partial_x$, and periodic boundary conditions. Finite differences introduce excessive dispersion and will usually be inadequate, although of theoretical interest to compare to known integrable systems formed with finite differences. We therefore need spectral methods. There are four choices: first, a full spectral method, multiplying out the nonlinear terms in $\mathcal{H}$ and truncating the introduced high frequencies; second, replacing ∂_x by the pseudo-spectral matrix D; and each method in Fourier or real space. The time-continuous dynamics are the same whether one works in Fourier or physical variables; usually the choice is made to minimize the number of Fourier transforms required per time-step. Details of particular spatial discretizations will be deferred to section 5.

2. Symplectic integration

Suppose the Hamiltonian may be split into two parts—the "P-Q" splitting:

$$H = T(p) + V(q).$$

Later we will also consider "L-N" splitting

$$H = L(u) + N(u)$$

where L has linear dynamics and N is nonlinear. Let $X_T = J\nabla T$, etc., be the associated vector fields, and e^{kX} be the time-k flow of the vector field X. If $J = \begin{bmatrix} 0 & K \\ -K^T & 0 \end{bmatrix}$ then the following map is an explicit, first-order approximation of the true flow e^{kX_H} (Ruth (1983)):

$$e^{kX_T} e^{kX_V} \equiv e^{kX} \tag{2.1}$$

which is computed as

$$p^{n+1} = p^n - k\left(K^T V'(q^n)\right), \qquad q^{n+1} = q^n + k\left(KT'(p^{n+1})\right).$$

In (2.1),

$$X = X_T + X_V + k[X_T, X_V] + \mathcal{O}(k^2) \equiv J\nabla\widehat{H} \tag{2.2}$$

and

$$\widehat{H} = H + H_0 + \mathcal{O}(k^2), \qquad H_0 = -k\{T, V\} \tag{2.3}$$

where the error (due to the noncommutativity of T and V) is expanded using the Campbell-Baker-Hausdorff (BCH) formula, and (2.3) follows because $[X_T, X_V] = -X_{\{T,V\}}$ (Arnol′d (1978)). The series in (2.2) is only an asymptotic series in k—it does not generally converge, although it will for linear systems for small enough k. A Hamiltonian whose flow is strictly equal to e^{kX} must be nonautonomous in general (MacKay (1992), McLachlan and Atela (1992)). Despite this problem, one might call H_0 the "autonomous Hamiltonian truncation error" (Yoshida (1991)).

The leapfrog ("LF2") method extends the method (2.1) to second order:

$$\varphi(k) = e^{\frac{1}{2}kX_T} e^{kX_V} e^{\frac{1}{2}kX_T} \equiv e^{kX_{\widehat{H}}},$$

$$\widehat{H} = H + \frac{k^2}{24}\left(2\{V,\{T,V\}\} - \{T,\{T,V\}\}\right) + \mathcal{O}(k^4) \tag{2.4}$$

It is *symmetric*, that is, $\varphi(k)\varphi(-k) = 1$. Suzuki (1991) and Yoshida (1990) use this property to construct schemes of arbitrary order by composing $(2s+1)$ leapfrog stages and preserving the symmetry:

$$\varphi(w_s k)\ldots\varphi(w_1 k)\varphi(w_0 k)\varphi(w_1 k)\ldots\varphi(w_s k). \tag{2.5}$$

where $w_0 = 1 - 2(w_1 + \ldots + w_s)$. Particular schemes are given in Table 1. A fourth-order scheme which has been rediscovered many times (Candy and Rozmus (1991), Suzuki (1991), Yoshida (1990)) is LF4a, which has $s = 1$ (see Table 1). However, this method takes a large backwards step of $1.70k$, leading to poor accuracy and stability. A better fourth order method LF4b (Suzuki (1991)), whose largest step is $-0.66k$, has $s = 2$, and one can show that this is close to the most accurate fourth-order method of this type. (All such methods of higher than second order must take a backwards step, Suzuki (1991).) The best sixth-order method, LF6a, is Yoshida's Method A which has $s = 3$.

Table 1

If $T(p)$ is quadratic (i.e. one may write $\ddot{q} = f(q)$), one can do significantly better by simply composing several stages of (2.1): $\varphi = \prod_{i=s}^{1} e^{a_i k X_T} e^{b_i k X_V}$. This is an example of a partitioned Runge-Kutta method, equivalent to a Runge-Kutta-Nyström (RKN) method if p is eliminated. The most accurate 4th- and 5th-order methods (in the sense of minimizing the Hamiltonian truncation error at constant work) are due to McLachlan and Atela (1992);

the 4th-order one, LF4c, has $s = 4$ stages. Okunbor and Skeel (1992) give sixteen 8-stage, 6th-order methods. Although symmetric, they are not formed by composing leapfrog steps, which allows more freedom in the choice of the a_i's and b_i's. Their method 13 (LF6b) has the smallest truncation error, about 0.02 times that of LF6a.

If both X_L and the nonlinear vector field X_N can be integrated exactly, then one may use the same composition methods with L-N splitting ($\varphi = e^{kX_L}e^{kX_N}$, etc.) This will usually be superior in that more (or all) of the derivatives in H will be treated exactly, and for weak nonlinearities, the truncation error will be much smaller. Such is the case for both NLW and NLS, which we consider below. Furthermore, the more accurate RKN methods may still sometimes be used:

Proposition 1: *Let* $J = \begin{bmatrix} 0 & K \\ -K^T & 0 \end{bmatrix}$*, where* K *is constant, and* $H = L(q,p) + N(q)$*, where* L *is a quadratic polynomial in* p *and* q*. Then any explicit canonical Runge-Kutta-Nyström method of order five or less, or any symmetric method of order six or less, maintains its order of accuracy when applied to this splitting.*

Proof: The special requirement of RKN for the splitting $H = T(p) + V(q)$ is that certain terms in the expansion of $e^{a_i k X_T} e^{b_i k X_V} \ldots$ contain a factor T''' and hence vanish identically. We compare these terms to those appearing in the expansion via the BCH formula, namely higher-order commutators ($\equiv$ Poisson brackets, see (2.2,2.3)) of

$$\{T, V\} = -W_j T_j, \qquad T_j = \frac{\partial T}{\partial p_j}, \quad W_j = \sum_i \frac{\partial V}{\partial q_i} K_{ij}.$$

The first vanishing term is at $\mathcal{O}(k^4)$, which corresponds to a 4th-order method. It is $\{V, \{V, \{T, V\}\}\} = T_{ijk} W_i W_j W_k$. If now the same method is applied to L-N splitting this term is $L_{ijk} W_i W_j W_k$ which is also identically zero (although $\partial L/\partial q$ does enter in the other terms). At fifth order, the two zero terms (in P-Q) are the two commutators of this one with T and V, hence also zero in L-N. At sixth order this simplicity breaks down: the twelve terms in the BCH expansion reduce to eight for both the P-Q and the L-N splitting. These eight contain five distinct terms in the P-Q case but eleven in the L-N case (the extra terms containing $\partial L/\partial q$, etc); hence the order conditions in the two cases are different. But if the method is symmetric, the sixth-order terms are identically zero (Yoshida (1990)), so the RKN methods do then carry over to the L-N splitting. ∎

(Note: Because the theorem is not true for all orders, a proof must involve looking at the error expansion; but we do not need the entire BCH series, just the knowledge of which terms can appear at each order, namely, the commutators of T and V (or L and N) of that order.)

Integrals of the system are conserved if they are integrals of each part of the Hamiltonian separately. This is clearly the case for linear integrals (conserved by any consistent scheme anyway) and for bilinear integrals under both the P-Q and L-N splittings. When J is constant, Casimirs are linear functions and hence conserved.

Example 2: De Frutos, Ortega and Sanz-Serna have given two treatments of the Boussinesq equation $\mathcal{J} = \begin{bmatrix} 0 & \partial_x \\ \partial_x & 0 \end{bmatrix}$, $\mathcal{H} = \int \left(\frac{1}{2}(p^2 + (q_x)^2 + q^2) + \frac{1}{3}q^3\right)\, dx$. The first method (de Frutos *et al.* (1991)) is unconditionally stable; this is achieved by time-averaging the stiffest ($D^4 q$) term:

$$(q^{n+1} - 2q^n + q^{n-1})/k^2 = -\frac{1}{4}D^4(q^{n+1} + 2q^n + q^{n-1}) + Dq^n + D((q^n)^2) \tag{2.6}$$

where superscripts denote time-levels and D is the pseudo-spectral difference operator, but could just as well be the (diagonal) spectral difference operator or even ∂_x. Rearranging terms, this can be written as a map φ:

$$\begin{aligned} p^{n+\frac{1}{2}} &= p^{n-\frac{1}{2}} + k(I + \frac{k^2}{4}D^4)^{-1}(-D^3 q^n + Dq^n + D(q^n)^2) \\ &\equiv p^{n-\frac{1}{2}} + k \qquad E \qquad\quad N(q^n) \\ q^{n+1} &= q^n + kDp^{n+\frac{1}{2}} \end{aligned} \tag{2.7}$$

showing that stability is achieved by braking the high modes severely—the temporal frequency ω of a spatial wavenumber m is reduced to $\omega/(1 + k^2 m^4/4)$ through the term E. In fact the highest frequency wave has wavenumber $1/\sqrt{2k}$, showing that $k/h^2 \sim 1$ is in fact required to capture a positive fraction of the waves as $h \to 0$. Secondly, a direct calculation of $\varphi' J \varphi'^T$ shows that φ is a Poisson map iff $EN' = N'E$, which is not true here. It is true if $N(q)$ is linear; in this case the method is equivalent to leapfrog with the high modes braked in the Hamiltonian. (A similar method of gaining unconditional stability is used in Dai (1992) for the variable-coefficient Schrödinger equation.)

Their second method (de Frutos *et al.* (1990)) is equivalent to (2.7) with $E = I$. In our framework this is P-Q splitting with time-stepping $e^{kX_T}e^{kX_V}$, which is second-order if the unknowns are staggered in time (q^n, $p^{n-\frac{1}{2}}$). They prove convergence and nonlinear stability for this method. Consider instead LF2, which is equivalent if the initial data for the staggered method are given by the mapping $e^{-kX_V/2}$ (i.e., $p_0^{-\frac{1}{2}} = p_0 - \frac{1}{2}kN(q_0)$). The Hamiltonian truncation errors (2.4) for the P-Q and L-N splittings are

$$\begin{aligned} &\text{P-Q:} \qquad -\frac{k^2}{24}\int -pp_{xx} + p^2 + 2qp^2 + 2(-q_{xx} + q + q^2)^2\, dx \\ &\text{L-N:} \qquad -\frac{k^2}{24}\int \qquad\qquad\qquad 2qp^2 \qquad\qquad\qquad + 2q^4\, dx \end{aligned}$$

or their corresponding discretizations. For strong nonlinearities, both are $\mathcal{O}(q^4)$; for $q \sim p \sim 1$, L-N has two terms against nine, and no derivatives (which can be larger); and for weak nonlinearities ($q \sim p \ll 1$), L-N is $\mathcal{O}(q^3)$ whereas P-Q is $\mathcal{O}(q^2)$. In addition, one may use the optimal RKN integrators; and the L-N splitting gains a factor $\frac{\pi}{2}$ in the stability criterion. ■

Example 3: Glassey (1992) considers the one-dimensional Zakharov equations

$$i\dot{E} + E_{xx} = qE \tag{2.8a}$$

$$\ddot{q} - q_{xx} = (|E|^2)_{xx} \tag{2.8b}$$

where E is complex (compare to NLS (1.2)) and proves nonlinear stability of a tridiagonally-implicit second-order scheme for the Zakharov equations—Crank-Nicolson on (2.8a) (equivalent to the midpoint rule, see below) and

$$(q^{n+1} - 2q^n + q^{n-1})/k^2 - D^2(q^{n+1} + q^{n-1})/2 = D^2(|E^n|^2)$$

on (2.8b), where D^2 represents central differencing of ∂_{xx}. (This method conserves a discrete energy but is not symplectic—see the discussion of analogous methods for NLW, section 5). But the L-N splitting is symplectic, explicit, has much smaller truncation errors and extends easily to any order: Let

$$\mathcal{J} = \begin{bmatrix} 0 & \partial_x & 0 \\ \partial_x & 0 & 0 \\ 0 & 0 & -i \end{bmatrix}, \qquad \mathcal{H} = \mathcal{L} + \mathcal{N},$$

$$\mathcal{L} = \int |E_x|^2 + \frac{1}{2}(q^2 + p^2)\, dx, \qquad \mathcal{N} = \int q|E|^2\, dx$$

(the spatial truncation is straightforward). Then e^{kX_L} is given by trigonometric functions, and e^{kX_N} can also be solved exactly:

$$\left.\begin{aligned} \dot{q} &= 0 \\ \dot{p} &= (|E|^2)_x \\ i\dot{E} &= qE \end{aligned}\right\} \quad \Rightarrow \quad \left\{\begin{aligned} q^{n+1} &= q^n \\ p^{n+1} &= p^n + k(|E^n|^2)_x \\ E^{n+1} &= \exp(-iq^n k)E^n \end{aligned}\right.$$

■

If e^{kX_N} cannot be found, or if H is not separable into parts which can be conveniently integrated, then an explicit symplectic integrator is usually not available. Instead one turns to the midpoint rule M2 (see Table 1). A generalization of M2 is that the Gaussian Runge-Kutta (GRK) methods (which have s intermediate unknowns and order $2s$) are also symplectic. Furthermore, they are also symplectic with respect to any constant Poisson matrix J, of any rank (MacKay (1992)). Thus they are suitable for discretizations of $\mathcal{J} = \partial_x$. Finally, they preserve quadratic integrals of the system, and preserve reversibility with respect to linear involutions (i.e. $S\mathrm{M2}(-k)S\mathrm{M2}(k) = 1$, where $S^2 = 1$) and hence are symmetric (the case $S = 1$). Therefore the above Yoshida-Suzuki methods all work if we replace a single leapfrog stage by second-order Gaussian Runge-Kutta, which is just the midpoint rule. This gives the methods M4a, M4b, and M6a. (M4a was independently discovered by Sanz-Serna and Abia (1991).) Ivanauskas (1991) proves convergence and stability with respect to initial conditions of the midpoint rule M2 applied to nonlinear Schrödinger equations.

3. Behavior of P-Q splitting

With finite differences, P-Q splitting would normally be used, because computing e^{kX_L} requires a Fourier transform. Even with L-N splitting, there is still the question of how the splitting acts on any derivatives remaining in N. For a straightforward analysis of this, we consider the linear wave equation $\dot q = p$, $\dot p = q_{xx}$ with P-Q splitting, and investigate the above methods with regard to *stability*, *dispersion* and *distortion*. The time-stepping is identical if one works in real or in Fourier space; choose the latter, so that the modes uncouple, and a change of scale reduces each to a linear oscillator:

$$\dot q = p, \qquad \dot p = -q.$$

Analysis of this system will apply to any linear Hamiltonian PDE with a nondegenerate elliptic fixed point (i.e. all eigenvalues are on the imaginary axis and all Jordan blocks of the Jacobian are 1×1). Write one time step of the method as an explicit linear map

$$\begin{pmatrix} q^1 \\ p^1 \end{pmatrix} = A(k)\begin{pmatrix} q^0 \\ p^0 \end{pmatrix} = \begin{pmatrix} A_{11}(k) & A_{12}(k) \\ A_{21}(k) & A_{22}(k) \end{pmatrix}\begin{pmatrix} q^0 \\ p^0 \end{pmatrix}$$

For the basic Ruth step $e^{akX_T}e^{bkX_V}$ we have

$$A = \begin{pmatrix} 1-abk^2 & ak \\ -bk & 1 \end{pmatrix}$$

and for leapfrog LF2,

$$A = \begin{pmatrix} 1-k^2/2 & k(1-k^2/4) \\ -k & 1-k^2/2 \end{pmatrix}.$$

For the midpoint rule, one solves a 2×2 system to find

$$A = \begin{pmatrix} 1-k^2/4 & k \\ -k & 1-k^2/4 \end{pmatrix}\frac{1}{1+k^2/4}.$$

For the higher-order Gaussian Runge-Kutta's one must solve larger systems. For GRK4, one finds

$$A_{11} = A_{22} = \frac{144-60k^2+k^4}{144+12k^2+k^4} \quad \text{and} \quad A_{12} = -A_{21} = \frac{12k(12-k^2)}{144+12k^2+k^4}.$$

For GRK6,

$$A_{11} = A_{22} = \frac{14400-6480k^2+264k^4-k^6}{14400+720k^2+24k^4+k^6}$$

and

$$A_{12} = -A_{21} = \frac{24k(600-70k^2+k^4)}{14400+720k^2+24k^4+k^6}.$$

For methods formed by composition, A is a product of such matrices (which we do not write out here), and we have by induction in the number of stages that A_{11} and A_{22} are even polynomials in k and A_{12} and A_{21} are odd. For symmetric methods, writing out the symmetry condition shows that $A_{11} = A_{22}$. Because the methods are symplectic, $\det A = 1$.

The exact solution for the linear oscillator is

$$A_0 = \begin{pmatrix} \cos k & \sin k \\ -\sin k & \cos k \end{pmatrix}$$

The matrix for a method of order p will agree with this up to terms of order k^p; thus the first wrong term in $\mathrm{tr}A$ is of order k^{p+2} for even-order methods. This error is available immediately for the explicit methods and can be found by Taylor series for implicit methods.

The method is stable if solutions remain bounded for all time. The eigenvalues of A are

$$\begin{aligned}\lambda &= \frac{\mathrm{tr}A}{2} \pm i\sqrt{1-\left(\frac{\mathrm{tr}A}{2}\right)^2} \\ &= \exp(\pm i\theta) \text{ where } \cos(\theta) = \frac{\mathrm{tr}A}{2}.\end{aligned}$$

Thus we have stability for a particular k iff $|\mathrm{tr}A(k)| \le 2$. Figure 1 shows $\cos(\theta)$ for each method, in comparison to the exact value $\cos(k)$; this shows the *accuracy* of each method for small k, and the *stability* for large k. This result can now be adapted to the original system:

Figs. 1 & 2

Proposition 2: *Consider $\ddot{q} = q_{xx}$ discretized with time-step k, spatial mesh size h, periodic boundary conditions and a symplectic integrator with matrix $A(k)$ defined above. Let k^* be the least positive root of $|\mathrm{tr}A(k)| = 2$. Then the stability criterion, depending on the spatial discretization, is*

(a) Pseudo-spectral differences: $\frac{k}{h} \le \frac{1}{\pi}k^*$

(b) Second-order finite differences: $\frac{k}{h} \le \frac{1}{2}k^*$

(c) Fourth-order finite differences: $\frac{k}{h} \le \frac{\sqrt{3}}{4}k^*$

Proof: For the spectral discretization with Fourier modes $-\frac{M}{2}+1 \le m \le \frac{M}{2}$, we require stability for each oscillator ($\ddot{q}_m = -m^2 q_m$) separately; rescaling leads to $mk \le k^*$ for $0 \le m \le \frac{M}{2}$. Then $M = 2\pi/h$ gives (a). For finite differences, von Neumann stability analysis reduces to this case with $m^2 = s$, where s is the Fourier symbol of the finite difference: $s = 4\sin^2(\frac{mh}{2})/h^2$ for the central difference $h^{-2}[1,-2,1]u$, and $s = (4\sin^2\frac{mh}{2} + \frac{4}{3}\sin^4\frac{mh}{2})/h^2$ for the fourth-order difference $(12h^2)^{-1}[-1,16,-30,16,-1]u$. We require stability for all wavenumbers m, and thus for the maximum value of s: 4 and $\frac{16}{3}$ respectively. This gives (b) and (c). ∎

Example 4: For leapfrog, $\mathrm{tr}A/2 = 1 - k^2/2$ so $k^* = 2$. We therefore have stability in the spectral approximation for $k/h \le \frac{2}{\pi} \sim 0.6366$. LF4a is worse: we find $2k^{*2} = 12 - 6(w + w^2) + 3\sqrt{-8+2w+4w^2}$, where $w = \sqrt[3]{2}$; $k^* \sim 1.5734$ and we need $k/h \le 0.5008$. Gaussian Runge-Kutta is unconditionally stable, which this example confirms because $|A_{11}| < 1$. ∎

Table 2

For the other methods the roots of the polynomials must be found numerically, and the corresponding stability criteria are given in Table 2. Notice that they are quite good—for non-symplectic methods (e.g. three-time-level leapfrog), one typically needs Courant numbers k/h near 1 with finite differences, and near $1/\pi$ with spectral differences. The results apply to any linear PDE with P-Q splitting: if the time-continuous problem has eigenvalues $i\sigma_m$, then the stability criterion is $k\sigma_m < k^*$.

No general time-integrator can be free of dispersion in general. Historically this has led to schemes which introduce artificial dissipation of the high modes to prevent "wiggles". Indeed Crank-Nicolson (equivalent to the midpoint rule for a linear PDE) is often frowned on for just this reason. Now we have expressly disallowed numerical dissipation. Does dispersion mean that we cannot expect good long-time behavior from symplectic integrators? Certainly it does in the case of the linear wave equation, for which any initial condition will eventually disperse into its constituent modes as all phase accuracy is lost. However, for nonlinear and particularly for near-integrable equations, we can hope that phase locking inherent in the system will prevent this. (Consider the ODE case of coupled oscillators, for example.) In addition, it turns out that some integrators (e.g. LF4c) and L-N splitting have negligible dispersion errors.

We take $k \le k^*$ and calculate the eigenvectors of A; separating real and imaginary parts show that the phase space is foliated by similar invariant ellipses, of which one is

$$B\begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix} \equiv \begin{pmatrix} A_{12} & 0\\ \cos\theta - A_{11} & \sin\theta\end{pmatrix}\begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix}.$$

Applying the map to this ellipse and using $\det A = 1$ gives

$$A\,B\begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix} = B\begin{pmatrix}\cos(\alpha-\theta)\\ \sin(\alpha-\theta)\end{pmatrix}.$$

Thus the map moves a point an angle θ around the ellipse each time-step, giving a dispersion relation θ. The exact map has $\theta = k$, and we are only considering a single wave, so the most natural error measure is the relative phase speed of that wave, $c = \theta/k$. Because $\cos\theta = \cos k + ak^{p+2} + o(k^{p+2})$ for even-order methods, we have $c \sim 1 - ak^p$ for small k (see Table 2). At $k = k^*$, $c = \pi/k^*$. Figures 1 and 2 show c for different methods. Despite being unconditionally stable, implicit methods are not superior here, and phase accuracy is lost completely if they are used with large values of k/h.

Finally, the invariant circles of the oscillator may become eccentric and tilted in the discrete system. We call this property *distortion.* If the ellipse has semi-major axis a and semi-minor axis b, then the energy $q^2 + p^2$ in this mode can change during the integration by up to a factor of $(a/b)^2$. So we define the distortion as $d = |(\frac{a}{b})^2 - 1|$. Because the ellipse can be rotated, this is tedious to calculate in general. However, for symmetric methods,

$\cos\theta = A_{11}$ so the matrix B is diagonal. So

$$\text{distortion } d = \left|\frac{(1-A_{11})^2}{A_{12}^2} - 1\right| = d_0 k^p + o(k^p)$$

and the leading term is shown in Table 2. Because GRK methods conserve the energy q^2+p^2, they have zero distortion.

4. Stability of L-N splitting

There are two approaches to the linear stability of splitting methods. Firstly, one can make general statements based on the generic bifurcations of symplectic vector fields and maps, giving sufficient conditions for linear stability. Secondly, the eigenvalues of the time-map can be computed explicitly for particular examples; the generic sufficient conditions turn out to be often necessary as well.

Here we are thinking of integrating $H = L+N$ with one of the composition methods in Table 1 with the resulting map $\widehat{\varphi}$ linearized about some steady state. Now $\widehat{\varphi} = e^{kX_{\widehat{H}}} = C$ is linear and hence is the time-k map of some autonomous linear Hamiltonian $\widehat{H}$ which can be found directly: $\widehat{H} = u^T B u$ where $C^T B C = B$. In this case the asymptotic series

$$\widehat{H} = H + k^p H_0 + \ldots \tag{4.1}$$

(cf. (2.4)) will converge to $\widehat{H}$ for k small enough. However, one cannot use the series to examine stability because near the onset of instability, typically all its terms are the same order in k. Examining the first term in (4.1) can determine when the conditions of the following proposition are *not* satisfied, and can help in choosing a good splitting of H. Then (roughly, if dispersion errors are $o(1)$ as $k\to 0$) two time-steps per period of the fastest wave are sufficient for stability:

Proposition 3: *Suppose X_H has pure imaginary eigenvalues $\{\pm i\sigma_m\}_{m=1}^M$, $\sigma_1 \le \ldots \le \sigma_M$, any multiple eigenvalues have positive signature, and any zero eigenvalues are associated with zeros of both X_L and X_N. Let $k\to 0$ with $k\sigma_M = k^*$ held fixed, and assume that in this limit $\widehat{\varphi}$ is a small perturbation of $\varphi = e^{kX_H}$. (One may need to rescale the independent variables to get $\varphi \sim 1$ first.) Then, for M sufficiently large, the method $\widehat{\varphi}$ is generically linearly stable for $k^* < \pi$.*

Proof: We are investigating the stability of the fixed point at the origin to small symplectic perturbations. The nonunit eigenvalues of φ are $e^{ik\sigma_m}$ which are are bounded away from -1 if $k^*<\pi$. Because the eigenvalues of the vector field have positive signature, so do those of its time-k flow. These are just the requirements for generic stability of the origin when φ is perturbed to $\widehat{\varphi}$ (Arnol′d (1978), MacKay (1986, 1992). Finally, if zero eigenvalues of

X_H come from zeros in X_L and X_N, then there is a corresponding zero in $X_{\widehat{H}}$, so the $+1$ eigenvalues of φ are fixed and do not split. ■

Notes:

1. If $H = T(p) + V(q)$ is separable, eigenvalues are guaranteed to have positive signature (MacKay (1986)).
2. At $\pm k\sigma_m = \pi$, φ has a double eigenvalue at -1. In the perturbed map $\widehat{\varphi}$ this generically splits into a real pair, signaling loss of stability in this mode. It may be a bubble of instability or a permanent loss.
3. If nonzero frequencies in L and N cancel out to give a zero eigenvalue in X_H, then this may indeed be falsely split by the integrator.
4. If the $\mathcal{O}(k^p)$ term in (4.1) is not subdominant to H then the above hypotheses cannot be satisfied. For example, take J canonical, $L = \frac{1}{2}(p^2 + m^a q^2)$ where m is the mode number, and $N = \frac{1}{2}m^b q^2$ (after linearizing around some steady state). That is, L is a linear wave equation with a derivatives of q, and N has b derivatives of q. Then the condition for subdominance is just $b < a$, i.e. the (originally nonlinear) term must have fewer derivatives than the linear term. Indeed, a direct calculation shows that the proposition applies to NLW (1.1), and to the Boussinesq equation (Example 2), with L-N splitting. For the linear wave equation with P-Q splitting, all terms are the same order, and in fact in the high modes, the perturbation to φ is $\mathcal{O}(1)$—hence the stricter stability conditions found in section 3.

Example 5: Nonlinear Wave Equations. Sine-Gordon $\ddot{q} - q_{xx} + \sin(q) = 0$ linearized about $q = 0$ is Klein-Gordon. Consider LF2 with L-N splitting on this equation: each mode decouples into a linear map

$$\begin{aligned} A &= e^{\frac{1}{2}kX_L} e^{kX_N} e^{\frac{1}{2}kX_L} \\ &= \begin{pmatrix} \cos(mk/2) & \sin(mk/2)/m \\ -m\sin(mk/2) & \cos(mk/2) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -k & 1 \end{pmatrix} \begin{pmatrix} \cos(mk/2) & \sin(mk/2)/m \\ -m\sin(mk/2) & \cos(mk/2) \end{pmatrix} \\ &= \begin{pmatrix} \cos(mk) - \frac{k}{2m}\sin(mk) & \frac{k}{2m^2}(\cos(mk)-1) + \sin(mk)/m \\ -\frac{k}{2}(\cos(mk)+1) - m\sin(mk) & \cos(mk) - \frac{k}{2m}\sin(mk) \end{pmatrix} \end{aligned}$$

$$\cos\theta = \frac{1}{2}\mathrm{tr}A = \cos(mk) - \frac{k}{2m}\sin(mk) \tag{4.2}$$

The stability limit for large m is indeed $mk < \pi$, but beyond this there are bubbles of instability. Consider k small with $mk = x$ held fixed, so $\frac{1}{2}\mathrm{tr}A = \cos x - \frac{1}{2}k^2\sin(x)/x$. Solving by series for x near $l\pi$ shows that this is larger than 1 in absolute value for $x \in (l\pi - k^2/l\pi, l\pi)$. However, these instabilities are unlikely to be triggered as they are only $k/l\pi$ times as wide as the spacing between modes, hence one can easily choose k so as to avoid them. The maximum growth rate in the bubbles is only $k^2/2\pi$.

Numerical experiments show that the formal stability limit can indeed be exceeded by a factor of three or four without the nonlinear terms triggering any instability, even for strong nonlinearities.

Expanding (4.2) for small k and any m gives the numerical dispersion relation:

$$\theta/k = \sqrt{1+m^2} + \frac{k^2}{24\sqrt{1+m^2}} + \mathcal{O}(k^4)$$

which is $\mathcal{O}(k^2)$ away from the true relation, uniformly in m. This, and the smaller truncation errors, are the great advantages of the L-N splitting. ■

5. Application 1: the Nonlinear Wave Equation

Consider the wave equation NLW

$$\dot{q} = p, \quad \dot{p} = -q_{xx} - V'(q) \tag{5.1}$$

on $[0, 2\pi]$ with periodic boundary conditions. After a Fourier transform $\mathcal{J}$ is canonical. (In fact it is multiplied by $1/2\pi$; we absorb this factor in $\mathcal{H}$.) The canonically conjugate variables are the Fourier coefficients q_m and p_{-m}. The Hamiltonian becomes

$$\mathcal{H} = \frac{1}{2}p_0^2 + \sum_{\substack{-\infty<m<\infty \\ m\neq 0}} \left(p_m p_{-m} + m^2 q_m q_{-m}\right) + \frac{1}{2\pi}\int_0^{2\pi} V\Big(\sum_{m=-\infty}^{\infty} q_m e^{imx}\Big)\, dx$$

The discretization consists in dropping all but a finite number of modes; the integral is replaced by the trapezoidal rule; and, with F representing the discrete Fourier transform, we get

$$H = \frac{1}{2}p_0^2 + \sum_{\substack{-M/2+1\leq m\leq M/2 \\ m\neq 0}} \left(p_m p_{-m} + m^2 q_m q_{-m}\right) + \frac{1}{M}\sum_{j=0}^{M-1} V((F^{-1}q)_j)$$

with Hamilton's equations

$$\begin{aligned} \dot{q}_m &= p_m \\ \dot{p}_m &= -m^2 q_m - (FV'(F^{-1}q))_m \end{aligned}$$

With the restriction that $q(x)$ be real, we can identify q_m and q^*_{-m}, getting an M degree-of-freedom Hamiltonian system in canonical form. This is equivalent to the standard pseudo-spectral truncation, but makes apparent the aliasing errors in the nonlinear term. These may be removed to any desired degree within the Hamiltonian formulation: take $M^+ > M$, let *pad* zero-extend a vector to length M^+, and *chop* truncate the $M^+ - M$ high frequencies. The nonlinear term in H becomes $\frac{1}{M^+}\sum_{j=0}^{M^+-1} V\left(F^{-1}\mathit{pad}(q)_j\right)$ with derivative $\mathit{chop}FV'\left(F^{-1}\mathit{pad}(q)\right)$. Quadratic nonlinearities will be completely anti-aliased with

$M/M^+ = \frac{2}{3}$ (the "two-thirds rule"); higher powers (q^s) by $M/M^+ = \frac{2}{s+1}$ (if we also drop the "odd man out" mode $m = M/2$ when s is odd). Complete anti-aliasing is equivalent to the full spectral method in which nonlinearities are multiplied out; but for non-power terms (e.g. $\sin u$) this is not practical, so it's better to stick to the above treatment.

We take $V(q) = q^4/4$ for some numerical tests. This equation is discussed at length in Strauss (1989): if the initial data are C^∞ then the solution stays C^∞ for all time; it also is known to have spatially odd, time-periodic solutions. (These may be easily computed by finding fixed points of a Poincaré map.) We ran numerical tests and confirmed the following properties of symplectic integrators on this equation: (a) the energy error did not increase secularly with time; (b) with initial conditions near an elliptic periodic orbit, the orbit did not drift away from it or towards it over extremely long times (many thousand periods); (c) the bilinear momentum integral $\int pq_x\,dx$ ($\sum mp_m q_{-m}$ for the ODE's) was conserved within round-off error; and (d) the above accuracy and stability analyses were confirmed (nonlinear terms could destabilize the calculation only when the solution was extremely poorly resolved spatially).

Next we compare the L-N and P-Q splittings. Throughout we use LF4c, the optimal fourth-order integrator, so that both methods use the same amount of work. For small $|q|$ (i.e. weak nonlinearity) the relative truncation error of the P-Q scheme (H_0/H) is $\mathcal{O}(1)$, against $\mathcal{O}(q^2)$ for L-N, so there is no comparison in this case. Fig 3 shows the maximum energy error for the two schemes with initial conditions $p^0 = 0$, $q^0 = a\cos x$ and $M = 32$ for two different amplitudes a. (The last mode q_{16} is dropped). The energy error did not grow with time for any stable k. The P-Q splitting was stable for $c = k/h <\sim 0.94$ (cf. Table 2) and the L-N for $c <\sim 14$ at $a = 0.5$ and for $c < 3$ at $a = 2$—L-N has no *linear* stability limit here. Clearly the L-N splitting is preferred wherever it is feasible.

Although anti-aliasing is often used in large-scale simulations of fluids, it is not clear why aliasing the high-frequency interactions is worse than simply setting them to zero, as the full-spectral method does. Perhaps the extra work of anti-aliasing would be better used to increase resolution? This is true for spatially resolved integrations, but not necessarily for marginally-resolved ones (such as most fluid problems). The following example, in which a full-spectral truncation converges much more regularly, illustrates this.

Figs. 3 & 4

Table 3

Solutions with initial conditions $p^0 = 0$, $q^0 = a\cos x$ are unstable to odd perturbations for $a >\sim 1.85$; this makes a sensitive test case. Figure 4 shows the complicated structure of the instability through 1500 periods of the primary $\cos x$ mode (t is in units of the fundamental period 2π). We took $q^0 = a\cos x + 10^{-12}\sin x$ and measured the growth of this instability for different spatial truncations (see Table 3; all results are converged in time). With full-spectral, even $M = 8$ gives good results; with pseudo-spectral, convergence is erratic, although when M is large enough both have similar accuracy. Also striking are the methods' completely different relative behavior for slightly different initial conditions.

In the literature there is an emphasis on deriving schemes which conserve a discrete analog of the energy exactly. Such schemes are generally not symplectic, due to the result of Ge (1988) on non-conservation of energy in symplectic integrators. Here we consider one such scheme, which does turn out to have good long-time properties—not because of its conserved quantity, but because it is *reversible*. Glassey and Schaeffer (1991) study the second-order implicit scheme

$$G(q^{n+1}, q^{n-1}, q^n) \equiv \frac{q^{n+1} - 2q^n + q^{n-1}}{k^2} - D^2 q^n + N(q) = 0 \tag{5.2}$$

where D^2 represents central differencing of q_{xx} and

$$N(q) = \frac{1}{4}(q^{n+1} + q^{n-1})((q^{n+1})^2 + (q^{n-1})^2)$$

which conserves a discrete energy but is not symplectic. (One may introduce $p^{n+\frac{1}{2}} = (q^{n+1} - q^n)/k$ to write this in our standard form). By comparison, LF2 with central differences and P-Q splitting can be written in the form (5.2) with $N(u) = (q^n)^3$ and is explicit with smaller truncation error.

There is a traveling wave solution $q(x,t) = g(x - ct) = g(\eta)$ where

$$\eta = {}_2F_1(\frac{1}{2}, \frac{1}{4}, \frac{5}{4}; \alpha^2 \beta g^4)\alpha g, \qquad \alpha = \frac{\pi \Gamma(\frac{3}{4})^2}{4\sqrt{2}\Gamma(\frac{5}{4})^2\sqrt{c^2-1}}, \quad \beta = \frac{1}{2(c^2-1)}.$$

This may be used to generate starting values for (5.2), and to check accuracy—phase errors are easily subtracted out by comparing Fourier amplitudes, giving the "shape error" $\| |q_m| - |q_m^{\mathrm{exact}}| \|$. The results are that both schemes (LF2 and (5.2)) show similar qualitative behavior: the position of the traveling wave is wrong by $\mathcal{O}(tk^2)$ but errors in the shape of the wave do not grow with time. LF2 is more accurate due to the extra $\mathcal{O}(k^2 q^3)$ truncation error in (5.2). A numerical calculation with $c = 1.5$ showed LF2 to be $13.5\times$ more accurate in shape and $17.5\times$ in wave speed. However, note that any difference equation of the form (5.2) is reversible if $G(q^{n+1}, q^{n-1}, q^n) = G(q^{n-1}, q^{n+1}, q^n)$, as here. We observed the same qualitatively good results for other reversible discretizations, energy-preserving or not (e.g., for $N(u) = \frac{1}{2}(q^{n+1} + q^{n-1})(q^n)^2$). By contrast, the introduction of any (even $\mathcal{O}(k^4)$) asymmetry in (5.2) causes the shape error to grow linearly with time.

Reversible systems inherit many of the properties of Hamiltonian systems near their symmetry plane (here, $p = 0$), for example, they have a KAM theorem, and eigenvalues of symmetric fixed points have the same restrictions as do fixed points of Hamiltonian systems (Roberts *et al.* (1991)). However, away from the symmetry plane they may have attractors and repellers, which are undesirable in a numerical method for a Hamiltonian system. (It is possible that (5.2), while not symplectic, could be made symplectic by a change of variables.

Equivalently, it might preserve a Poisson bracket, but the wrong one. While not as attractive as preserving the correct one, the method would then have many of the same properties as standard symplectic integrators. This is unlikely but we have not been able to rule it out.)

Although the traveling wave is not symmetric, its symmetric partner is just itself shifted by π, so the above eigenvalue results extend to this case. Therefore we believe that any good long-time behavior of (5.2) is due to its reversibility, not its energy-conservation.

6. Application 2: The nonlinear Schrödinger equation

Consider the focusing nonlinear Schrödinger equation,

$$i\dot{\psi} + \psi_{xx} + 2|\psi|^2\psi = 0 \tag{6.1}$$

with periodic boundary conditions $\psi(0,t) = \psi(L,t)$; the *symmetric* case has, in addition, $\psi(x,t) = \psi(L-x,t)$ enforced $\forall t$. The plane waves $\psi = e^{2i|a|^2t}a$, $a \in \mathbb{C}$, form a two-dimensional set of periodic orbits. These orbits can be unstable, though, and in the symmetric case have an $N_h = \lfloor |a|L/\pi \rfloor$ dimensional homoclinic manifold connecting periodic orbits with the same $|a|$ but different phase. Orbits with initial conditions close to the plane wave will trace out different pieces of this manifold, but Ablowitz and Herbst (1990, and other references therein), using integrable and nonintegrable discretizations, have found these orbits difficult to calculate numerically. The problems became progressively more severe as N_h increases. For the standard central difference approximation (the "diagonal discretization") and adaptive Runge-Kutta-Merson time-integration they observe

a) a rapid flow of energy into the high frequencies ('spatial chaos'), or

b) temporal chaos in the time-series $\psi(x_0,t)$, or

c) for the unsymmetric problem, sudden loss of symmetry and increase in temporal chaos.

They emphasized that an integrable discretization, with its own (perturbed) homoclinic structure, was necessary to integrate NLS, and a comparison of the diagonal discretization with a second-order integrable discretization showed that the latter avoided problem (a) and was generally more reliable. However, for larger N_h ($>\sim 4$) it too faced the same problems. Some of these have been analyzed in terms of the inverse-scattering eigenvalues of the PDE. Instead, here, we emphasize the intrinsic numerical characteristics of the equations which must be dealt with for a successful integration. In our view integrability is not essential: the splitting of stable and unstable manifolds due to a nonintegrable discretization is generally exponentially small in the perturbation and is insignificant in a well-resolved calculation. A method that addresses the numerical characteristics of the *linear* equation $i\dot{\psi} + \psi_{xx} = 0$ will also cope with NLS. The following are the three crucial factors:

i) The homoclinic orbits have fine spatial structure—at least $40N_h$ grid points are needed just to represent them smoothly. Problem (a) is caused by insufficient spatial resolution;

the perturbation from the PDE is then no longer small and all accuracy is lost. The formation of 'wiggles' near sharp fronts is well known in both finite-difference and spectral calculations. In addition, here this is accompanied by breakup of the homoclinic structure. Forest *et al.* [10] saw similar behavior in a discretized Sine-Gordon, which showed a transition to chaos at the onset of instability in the PDE. From our point of view, it would be more correct to say that chaos is expected if the true solution is not smooth on the grid—the perturbation from integrability in then no longer small.

ii) The problem is numerically stiff because the linear part of (6.1) has eigenvalues $\sigma_m = i\mu_m^2$ (where $\mu_m = m\frac{2\pi}{L}$ and m is large), thus we expect to need $k\mu_m^2 < \pi$ just to resolve the fast waves temporally. Some numerical methods may allow this limit to be exceeded; one should not be surprised if this allows errors to build up in time.

iii) A symplectic integrator must be used for the time-integration. In particular, standard Runge-Kuttas are strongly dissipative below their stability limit (e.g. the standard fourth order RK damps waves by 0.5 per time-step at $k/h \sim 0.75$) which leads to substantial errors in the known integrals of the PDE.

We outline the numerical properties of a symplectic, pseudo-spectral scheme. We follow Ablowitz and Herbst (1990) and use $L = 4\sqrt{2}\pi$ throughout. L-N splitting works here: $e^{kX_L} = e^{ik\mu_m^2}\psi_m$ is performed in Fourier space, and $e^{kX_N} = e^{-2i|\psi_j|^2 k}\psi_j$ in real space. Anti-aliasing is not economic because the nonlinear part would need to be integrated in real space on the finer grid. Linearized around $\psi = 0$, the splitting gives the exact solution. Linearizing around the plane wave ψ_0, i.e. $\psi = \psi_0(1 - \varepsilon \cos \mu_m x)$ gives

$$\dot{\varepsilon} = i\mu_m^2 \varepsilon + 4i|a|^2 \Re\varepsilon \tag{6.2}$$

(Note: until the last paragraph in this section we are dealing with symmetric NLS.) Separating real and imaginary parts and applying LF2 gives (similarly to (4.2))

$$\frac{1}{2}\mathrm{tr}A = \cos(k\mu_m^2) + 2|a|^2 k \sin(k\mu_m^2)$$

and a series expansion shows that the time-discrete system has instability for $\mu_m < 2|a| - \frac{2}{3}|a|^5 k^2 + \mathcal{O}(k^4)$. The numerical growth rate, $\sigma_{num} = \ln(\lambda)/k$ where the eigenvalue $\lambda = \frac{1}{2}\mathrm{tr}A + \sqrt{(\frac{1}{2}\mathrm{tr}A)^2 - 1}$ is

$$\sigma_{num} = \mu_m\sqrt{4|a|^2 - \mu_m^2} - k^2 \frac{2|a|^4\mu_m^3}{3\sqrt{4|a|^2 - \mu_m^2}} + \mathcal{O}(k^4)$$

so we expect good results in the linear regime if $4|a|^2 - \mu_m^2$ is not small, i.e. if a mode is not marginally unstable.

However, as for sine-Gordon, there are bubbles of instability. Essentially the true, low-frequency instability aliases onto the high frequencies if fewer than two time-steps per period are taken. The first bubble is at $k\mu_m^2 \in (\pi, \pi + 4|a|^2 k)$. The frequency spacing near here is about $2k\mu_m$ (μ_m large) so the bubble may be avoided. However, in the full nonlinear system we do observe a weak instability in the marginally stable wave with $k\mu_m^2 = \pi$. Hence one should stick to $k\mu_m^2 < \pi$ for safety.

We use the second-order LF2 for illustration. In general, one expects no energy drift if the time-step is sufficient small; here, this means sufficiently small with respect to the period of the fastest waves. Indeed, although $k\mu_m^2 \sim \pi$ turns out to be sufficient for spatially well-resolved calculations, k needs to be smaller than this if there is less resolution—i.e. it is then necessary to integrate the high frequencies more accurately. A corollary is that near the stability limit, one should not expect a higher order method (e.g. LF4b) to be more accurate, and it is not. In fact higher-order methods are even more sensitive to k being too large. Secondly, the energy may be well conserved for a certain time and then jump. This can happen when the orbit enters a particularly steep region of phase space (i.e. all the energy concentrated in one unstable mode). Again, smaller time-steps ameliorate this problem. Thirdly, as an example of the sensitivity of this system, we found that in single precision it was impossible to stop the energy error increasing each time around the homoclinic manifold.

Figs. 5, 6, & 7

Another diagnostic is conservation of an integral of NLS,

$$I_5 = \int_0^L |\psi_{xx}|^2 + 2|\psi|^6 - 6|\psi_x|^2|\psi|^2 - ((|\psi|^2)_x)^2\, dx$$

which, after H, is the next-most complicated amplitude integral. (The first integral, $\int_0^L |\psi|^2\, dx$, is conserved exactly by LF2). The pseudo-spectral approximation of I_5 is not an integral of our ODE's, but we observe that it does not drift in a sufficiently-resolved (in time and space) calculation. As $k \to 0$, relative errors in I_5 are the same order as the last Fourier mode $|\psi_{M/2}|$. If this is large enough, or if the time-step is too large, I_5 may jump in regions of phase space where the solution is particularly steep. Figure 5 compares well- and marginally-resolved calculations for the case $N_h = 2$ (the initial condition is $\psi = a(1 - 0.1\cos x)$). If the above guidelines are followed, increasing N_h presents no further difficulties: Figure 6 shows results for $N_h = 8$. Note that we are not claiming that this shows the *exact* solution, only that it is qualitatively correct, e.g. the orbits stays near the homoclinic manifold and its temporal spectrum has rapid decay.

High *accuracy*, such as that of the pseudo-spectral method, is not crucial to a successful integration. Qualitatively similar results are obtained with the (nonintegrable, second-order finite difference) diagonal discretization with the same number of grid points. Hence we conclude that resolution (so that discretization errors are uniformly small) and symplectic time-integration are the essential factors here; accuracy and integrability are not as important.

The unsymmetric problem is more subtle, and we find that further resolution is needed to capture the dynamics even qualitatively. Problem (c) above is caused by the introduction of tiny asymmetry from round-off error which then grows catastrophically and soon becomes $\mathcal{O}(1)$. We have introduced a small asymmetry ($\psi = a(1-0.1\cos x)+10^{-8}\sin x$) and watched it evolve (see Figure 7.) Although it is supposed to grow, the problem is that in poorly-resolved regions of phase space, the asymmetry grows much too quickly (the sudden growth for $M = 64$ is triggered by a sharp spike in ψ). The reason is that in the unsymmetric problem, the eigenvalues of the linearized flow (6.2) become double, a property which is not structurally stable. If the evolution is sensitive to the false splitting of this eigenvalue pair, then one cannot hope to get good results without actually following the whole solution accurately. This applies to any non-structurally-stable flow. We note that recent work of Ablowitz *et al.* (1993) indicates that for even more unstable modes ($N_h \sim 8$), even spatially resolved calculations (with resolution that makes the spatial truncation errors as small as they are in the successful calculations above) show a spurious odd-mode instability, apparently related to round-off error.

7. Application 3: The Korteweg-de Vries equation

Lastly we consider an equation for which one must use implicit methods—KdV, (1.3). This been considered by de Frutos and Sanz-Serna (1991), who found that M4a could have better pointwise accuracy than (non-symplectic) leapfrog. The nonlinear vector field $\dot{u}_i = \sum_j D_{ij}u_j^2$ cannot be solved exactly so L-N splitting is not applicable. (Note that one can solve the corresponding continuous problem by characteristics, but this is of no use in deriving a symplectic map for the discrete problem). One might replace e^{kX_N} by an approximation (e.g. the midpoint rule) of the same order as the method, but this has no clear advantage in accuracy. Thus for fourth-order methods there are three choices, all implicit: GRK4, M4b, and M4a. In each case the linear part of the equations may be moved on to the left-hand side of the iteration; this requires solving a 2×2 system for GRK4. With $\dot{u} = L(u) + N(u)$ the iteration is

$$f_i^{(l+1)} = \frac{1}{1-\frac{1}{2}kz+\frac{1}{12}(kz)^2}\Big(c_i L(u) + N(u + k\sum_{j=1}^{2} a_{ij} f_j^{(l)})\Big)$$

where $c_1 = 1 - \frac{\sqrt{3}}{6}kz$, $c_2 = 1 + \frac{\sqrt{3}}{6}kz$, $A = (a_{ij}) = \begin{pmatrix} \frac{1}{4} & \frac{1}{4}-\frac{\sqrt{3}}{6} \\ \frac{1}{4}+\frac{\sqrt{3}}{6} & \frac{1}{4} \end{pmatrix}$, and z is an (imaginary) eigenvalue of the linear operator L; for KdV on $[0, L]$, $z = i(2\pi n/L)^3$. This iteration should be converged to round-off error to preserve symplecticity, and then the value of u at the new time-step is $u + k(f_1 + f_2)/2$. The two real FFTs needed to compute N here may be done as one complex FFT; in practice we found GRK4 to be only slightly slower to converge than the second-order midpoint rule M2.

Thus the work requirements per iteration for the three methods GRK4, M4b, and M4a are in the ratio 1 : 5 : 3; taking into account their different convergence rates (due to different internal time-steps) the total work was in the ratio 1 : 2.5 : 2.25; and for a one-soliton solution (see de Frutos and Sanz-Serna (1991)), the energy errors were in the ratio 1 : 10 : 100. (The difference in truncation errors is not this extreme; a possible explanation is the simple form of the nonlinearity in KdV). So the energy errors at constant work are in the ratio 1 : 390 : 2560. We conclude that GRK4 is the best method, although not as easy to program as M4b.

We saw in section 5 that for some equations and some initial conditions, reversible schemes are sufficient for good long-time behavior. With $\mathcal{J} = \partial_x$, the involution which reverses time is $x \to -x$. If one is using finite differences then the following type of scheme is $\mathcal{O}(kh, k^2, h^2)$ and reversible (Kruskal 1972, unpublished):

$$\frac{1}{k}(u_j^{n+1} - u_j^n) = \frac{1}{2h}(f_{j+1}^{n+1} - f_j^{n+1} + f_j^n - f_{j-1}^n)$$

where $u_j^n = u(nh, jk)$ and $f_j^n = \partial H/\partial u_j$ at time level n. Its advantage over M2 is that, although implicit, it has rapid convergence when swept from left to right. It can be made symmetric (and $\mathcal{O}(k^2, h^2)$) by adding a second stage with x reversed, which would be swept from right to left. The order could now be increased to four by composing several steps (as in (2.5)). However, because still four sweeps per (symmetric) step would be needed, this approach has no clear advantage over midpoint-based methods, considering that symplecticity has been abandoned.

Acknowledgements. I thank Jim Curry and Harvey Segur for their helpful comments on an early version of this paper, and particularly Mark Ablowitz and Connie Schober for many useful discussions on the material in section 6. The application of L-N splitting to the nonlinear Schrödinger equation was suggested by Connie Schober. The paper was improved thanks to the helpful comments of the referees.

References

Ablowitz, M., and Herbst, B.M. (1990): On homoclinic structure and numerically induced chaos for the nonlinear Schrödinger equation, SIAM J. Appl. Math. **50**(2) 339–351

Ablowitz, M.J., Schober, C., and Herbst, B.M. (1993): Numerical chaos, roundoff error, and unstable wave processes, submitted to Phys. Rev. Lett.

Arnol′d, V. (1978): Mathematical Methods of Classical Mechanics, Springer-Verlag, New York

Candy, J. and Rozmus, W. (1991): A symplectic integration algorithm for separable hamiltonian functions, J. Comp. Phys. **92**, 230–256.

Dai, Weizhong (1992): An unconditionally stable three-level explicit difference scheme for the Schrodinger equation with a variable coefficient, SIAM J. Num. Anal. **29**(1) p. 174

de Frutos, J., Ortega, T., and Sanz-Serna, J.M. (1991): Pseudospectral method for the "good" Boussinesq equation, Math. Comp. **57**(195) 109–122

de Frutos, J., Ortega, T., and Sanz-Serna, J.M. (1990): A Hamiltonian, explicit algorithm with spectral accuracy for the "good" Boussinesq equation, Comp. Meth. Appl. Mech. Eng. **80** 417–423

de Frutos, J., and Sanz-Serna, J.M. (1991): An easily implementable fourth-order method for the time-integration of wave problems, Report 1991/2, Departamento de Matemática Aplicada y Computación, Universidad de Valladolid, Valladolid, Spain

Fei, Zhang and Vasquez, Luis (1991): Two energy-conserving numerical schemes for the Sine-Gordon equation, App. Math. Comp. **45**(1) 17–30

Feng Kang and Qin Meng-Zhao (1987): The symplectic methods for the computation of Hamiltonian equations, *Numerical methods for partial differential equations, Proc. Shanghai* 1986, Lecture Notes in Mathematics **1297** 1–37, ed. Zhu You-lan and Gu Ben-yu, Springer, Berlin.

Forest, M.G., Goedde, C.G., and Sinha, A. (1992): Instability-driven energy transport in near-integrable, many degree-of-freedom, Hamiltonian systems, Phys. Rev. Lett. **68**(18), 2722–2724

Ge Zhong and Marsden, J.E. (1988): Lie-Poisson Hamilton-Jacobi theory and Lie-Poisson integrators, Phys. Lett. A **133**, 134–139.

Glassey, R.T. (1992): Convergence of an energy-preserving scheme for the Zakharov equations in one space dimension, Math. Comp. **58**(197), 83–102

Glassey, R.T. and Schaeffer, J. (1991): Convergence of a second-order scheme for semilinear hyperbolic equations in $2+1$ dimensions, Math. Comp. **56**(193) p. 87

Ivanauskas, F.F. (1991): Difference schemes for nonlinear equations of Schrödinger type, Sov. Math. Doklady **42**(2) 287–291

MacKay, R.S. (1992): Some aspects of the dynamics and numerics of Hamiltonian systems, in The dynamics of numerics and the numerics of dynamics, eds. D.S. Broomhead and A. Iserles (Oxford)

MacKay, R.S. (1986): Stability of equilibria of Hamiltonian systems, Nonlinear phenomena and chaos (ed. Sarben Sarkar) 254–270

McLachlan, R.I., and Atela, P. (1992): The accuracy of symplectic integrators, Nonlinearity **5** 541–562

Okunbor, D.I., and Skeel, R.D. (1992): Canonical Runge-Kutta-Nyström methods of orders 5 and 6, sub. to Math. Comp.

Olver, P.J. (1986): Applications of Lie groups to differential equations, New York: Springer-Verlag

Olver, P.J. (1988): Darboux' theorem for Hamiltonian differential operators, J. Diff. Eq. **71**(1), 10–33

Roberts, J.A.G., and Quispel, G.R.W. (1991): Chaos and time-reversal symmetry: order and chaos in reversible dynamical systems, Physics Report 91–32, Instituut voor Theoretische Fysica, Universiteit van Amsterdam, Valckenierstraat 65, 1018 XE Amsterdam, The Netherlands

Ruth, R.D. (1983): A canonical integration technique, IEEE Trans. Nucl. Sci. NS-**30**, 2669–2671.

Sanz-Serna, J.M. and Abia, L. (1991): Order conditions for Canonical Runge-Kutta schemes, SIAM J. Numer. Anal. **28**(4), pp. 1081–1096.

Strauss, W.A. (1989): Nonlinear wave equations, AMS Regional Conference Series no. 73

Suzuki, M. (1991): General theory of fractal path integrals with applications to many-body theories and statistical physics, J. Math. Phys. **32**(2), 400–407

Yoshida, H. (1990): Construction of higher order symplectic integrators, Phys. Lett. A **150**, 262–269.

Yoshida, H. (1991): Conserved quantities of symplectic integrators for Hamiltonian systems, preprint