Conservative Discretization Methods for Wave Equations, Lattice Equations, and Conservation Laws

Brenton LeMesurier
Department of Mathematics, College of Charleston, South Carolina

Escuela de Verano
Ecuaciones de Ondas Dispersivas
IIMAS, UNAM

Junio 4-8, 2012

Abstract

A variety of wave phenomena arise in systems from classical mechanics, quantum mechanics, nonlinear optics and molecular modeling which are modeled with mildly nonlinear differential equations of Hamiltonian form, with several invariant quantities and related invariance of the equations and solutions under transformation groups.

There is much evidence and theory that numerical solutions often respect the qualitative features of solutions better when one uses discretizations that respect these invariance features of the differential equations.

In these lectures, I will review a variety of approaches to conservative discretization of Hamiltonian systems of ODE's, with an emphasis on methods applicable to large quasilinear systems of ODEs, such as arise from spatial discretization of partial differential equations, and in models of molecular chains or lattices.

The results are different in some details than for smaller, more highly nonlinear systems, such as modeling planetary systems.

Contents

1 Prologue: Some Dispersive PDE's and Discrete Counterparts

1.1 The Equations

A few famous examples indicate the sort of nonlinear wave equations whose numerical solution will be considered in these lectures:

The Nonlinear Klein-Gordon equation [NLKG]

$$\frac{\partial^2 u}{\partial t^2} - \Delta u = g(u), \tag{1}$$

The Nonlinear Schrödinger Equation [NLS]

$$\mathrm{i}\frac{\partial z}{\partial t} + \Delta z + g(|z|^2)z = 0, \tag{2}$$

A Discrete Nonlinear Schrödinger Equation [DNLS]

$$\mathrm{i}\frac{dz_n}{dt} + \kappa(z_{n-1} - 2z_n + z_{n+1}) + g(|z_n|^2)z_n = 0. \tag{3}$$

We will also consider some less familiar examples.

First, one that introduces additional challenges to the design of a good numerical methods is

Davydov's Model of a Molecular Chain [DMC]

$$\mathrm{i}\frac{dz_n}{dt} + \kappa(z_{n-3} + z_{n+3}) - \lambda(z_{n-1} + z_{n+1}) = (q_{n+3} - q_{n-3})z_n, \tag{4}$$

$$m_0\frac{d^2 q_n}{dt^2} - (q_{n-3} - 2q_n + q_{n+3}) = |z_{n+3}|^2 - |z_{n-3}|^2, \tag{5}$$

[Davydov:1971, Davydov+Kislukha:1973]

More specifically, this describes α-helix protein, a structure that typically arises for short segments within protein molecules, in which the helix has roughly three amino acid residues in each twist, so that sequences of every third residue are aligned in roughly straight lines or *spines*, with strong attractive dipole coupling along these spines. For more general polypeptide molecules, a similar form could be used, but with position-dependent coefficients.

- The *exciton* variable z_n comes from Schrödinger's equation, and gives the probability that the C=N double bond at the n-th amino acid residue is in a rotational excited state, the only quantum mechanical excitation that is likely at biologically relevant temperatures.
- The mechanical variable q_n indicates the displacement of the n-th residue from its rest position.
- The λ terms relate to the repulsive interaction between excitons in residues that are adjacent along the molecular backbone.
- The κ terms relate to the attractive interaction between excitons in residues that are adjacent along spines.
- The remaining "mechanical" terms relate to the attractive electrostatic dipole interaction between resides adjacent along spines.

A Discrete Nonlinear Schrödinger Equation with a Non-local Nonlinearity

For both testing numerical methods and some theoretical studies, it is useful to simplify the Davydov model with two approximations:

- the fact that the signal speed of the mechanical coupling is far higher than for the exciton coupling, leading to the fast, small vibration limit, equivalent to $m_0 \to 0$, and
- the assumption of slow variation along the attractively coupled spines, leading to a phase shift of approximately $\pm 2\pi/3$ between adjoining residues and essentially the same solution on each spine, up to this phase shift.

The resulting equation for a single spline, with indexing changed to eliminate the steps of three, is

The Non-Local Discrete Schrödinger Equation [NLDNLS]

$$\mathrm{i}\frac{dz_n}{dt} + \kappa(z_{n+1} + z_{n-1}) + \left(\frac{1}{2}|z_{n-1}|^2 + |z_n|^2 + \frac{1}{2}|z_{n+1}|^2\right) z_n = 0. \tag{6}$$

This is an interesting challenge for numerical methods: it cannot be handled by some earlier approaches developed for discretizations of PDEs, due to the nonlinearity of non-local form, and it cannot be handled by some other popular methods for Lagrangian and mechanical systems to be discussed in these lectures.

Interactions in mono-molecular layers of molecules can be modeled by

A Two-dimensional Lattice Nonlinear Schrödinger Equation [LNLS]

$$\mathrm{i}\frac{dz_{m,n}}{dt} + (z_{m-1,n-1} + z_{m-1,n} + z_{m,n-1} + z_{m,n+1} + z_{m+1,n} + z_{m+1,n+1}) + g(|z_{m,n}|^2)z_{m,n} = 0. \tag{7}$$

The coupling is to the six nearest neighbors in a triangular lattice, with the indices corresponding to motion in directions at 60 degrees, not at a right-angle.

1.2 Some Key Features of the Equations

With appropriate boundary conditions (or endpoint conditions), these equations and others described below have some features that will be important for these lectures.

Conserved Energy, or Hamiltonian

Each equation has a conserved energy or *Hamiltonian* $\mathcal{H}$. (Here and from now on, $G' = g$, with $G(0) = 0$.)

$$\mathcal{H} = \int \frac{1}{2}\left(\frac{\partial u}{\partial t}\right)^2 + \frac{1}{2}|\nabla u|^2 - G(u)\, d\mathbf{x}, \qquad \text{or with } v = \frac{\partial u}{\partial t},$$

$$\mathcal{H} = \int \frac{1}{2}v^2 + \frac{1}{2}|\nabla u|^2 - G(u)\, d\mathbf{x} \qquad \text{for NLKG} \tag{8}$$

$$\mathcal{H} = \int |\nabla z|^2 - G(|z|^2)\, d\mathbf{x} \qquad \text{for NLS} \tag{9}$$

$$\mathcal{H} = \sum_n \kappa(z_{n+1} - z_n)(z_{n+1}^* - z_n^*) - G(z_n z_n^*) \qquad \text{for DNLS} \tag{10}$$

The Davydov system DMC has

$$\mathcal{H} = -\kappa \sum_n (z_n z^*_{n+3} + z_{n+3} z^*_n) + \lambda \sum_n (z_n z^*_{n+1} + z_{n+1} z^*_n) + \sum_n \left[\frac{p_n^2}{2m_0} + \frac{1}{2}(q_{n+3} - q_n)^2 \right] + \sum_n (q_{n+3} - q_{n-3}) z_n z^*_n. \tag{11}$$

while for the non-local discrete Schrödinger equation NLDNLS

$$\mathcal{H} = \sum_n -\kappa (z^*_n z_{n+1} + z_n z^*_{n+1}) + \frac{1}{2}(z^*_n z_n)(z^*_n z_n + z^*_{n+1} z_{n+1}), \tag{12}$$

and for the Lattice Nonlinear Schrödinger Equation LNLS

$$\mathcal{H} = \sum_{m.n} (z_{m,n} z^*_{m+1,n} + z_{m,n} z^*_{m,n+1} + z_{m,n} z^*_{m+1,n+1}) - \sum_n G(z_{n,n} z^*_{n,n}). \tag{13}$$

Hamiltonian Equations

Each equation has a Hamiltonian form, as will be explained more below.

$$\frac{\partial u}{\partial t} = \frac{\delta \mathcal{H}}{\delta v}, \quad \frac{\partial v}{\partial t} = -\frac{\delta \mathcal{H}}{\delta u} \qquad \text{for NLKG} \tag{14}$$

$$\frac{\partial z}{\partial t} = \mathrm{i} \frac{\delta \mathcal{H}}{\delta z^*} \qquad \text{for NLS} \tag{15}$$

$$\frac{d\mathbf{z}}{dt} = \mathrm{i} \frac{\partial \mathcal{H}}{\partial \mathbf{z}^*} \qquad \text{for DNLS and NLDNLS} \tag{16}$$

$$\frac{d\mathbf{z}}{dt} = \mathrm{i} \frac{\partial \mathcal{H}}{\partial \mathbf{z}^*} \quad \frac{d\mathbf{q}}{dt} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}}, \quad \frac{d\mathbf{p}}{dt} = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}} \qquad \text{for DMC} \tag{17}$$

Note: for the PDE's, this involves Fréchet functional derivatives in place of gradients.

Missing Lagrangian Forms, Non-Separable Hamiltonians

The Hamiltonian forms for the "Schrödinger-like equations" do not come from a Lagrangian form in the way seen with mechanical systems.

Also, they do not separate as a sum of kinetic energy $T(\mathbf{p})$ and potential energy $U(\mathbf{q})$, as happens with mechanical systems.

Conserved Momenta

Some of the equations also have a conserved momentum $\mathcal{P}$, quadratic (or linear) in the unknowns:

$$\mathcal{P} = \int \frac{\partial u}{\partial t} \nabla u \, d\mathbf{x} = \int v \nabla u \, d\mathbf{x} \qquad \text{for NLKG} \tag{18}$$

$$\mathcal{P} = \mathrm{i} \int z \nabla z^* - z^* \nabla z \, d\mathbf{x} \qquad \text{for NLS} \tag{19}$$

$$\mathcal{P} = \sum_n p_n \qquad \text{for DMC} \tag{20}$$

Conserved Charge

The equations of "Schrödinger type" have a conserved *charge* $\mathcal{E}$ (also called *exciton number* or *power* depending on the physical application). This is related to the probability density of quantum mechanics, and is again quadratic:

$$\mathcal{E} = \int |z|^2 \, d\mathbf{x} \qquad \text{for NLS} \tag{21}$$

$$\mathcal{E} = \sum_n z_n z_n^* \qquad \text{for DNLS, NLDNLS and DMC} \tag{22}$$

1.3 Some Notation

Vectors are indicated by bold face, as with $\mathbf{y}$ and $\mathbf{f}$.

We will use the notations $D_y f$ and $\dfrac{\partial f}{\partial y}$ flexibly, to indicate either individual partial derivatives

$$D_k f = D_{y_k} f = \frac{\partial f}{\partial y_k},$$

gradient vectors

$$D_{\mathbf{y}} f = \frac{\partial f}{\partial \mathbf{y}} = \left\langle \frac{\partial f}{\partial y_1} \quad \frac{\partial f}{\partial y_2} \quad \cdots \right\rangle,$$

or Jacobian matrices

$$D_{\mathbf{y}} \mathbf{f} = \frac{\partial \mathbf{f}}{\partial \mathbf{y}} = \left\{ \frac{\partial f_i}{\partial y_j} \right\}.$$

2 Review of Hamiltonian and Lagrangian Theory

2.1 Lagrangian Form and Hamilton's Principle

Lagrangian Form

The Lagrangian form of equations of motion is determined a *Lagrangian function* $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})$, where $\mathbf{q} = \{q_n\}_{n=1}^N$ is an N-component vector of generalized position coordinates, through the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_n}\right) = \frac{\partial \mathcal{L}}{\partial q_n}, \quad 1 \leq n \leq N. \tag{23}$$

We will mostly use the vector form

Euler-Lagrange Equations

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}\right) = \frac{\partial \mathcal{L}}{\partial \mathbf{q}}. \tag{24}$$

Hamilton's Principle

Hamilton's Principle says that the solution of the Euler-Lagrange equations minimizes the *action integral*

$$\mathcal{S}(\mathbf{q}) = \int_a^b \mathcal{L}(\mathbf{q}(t), \dot{\mathbf{q}}(t))\, dt$$

subject to the endpoint conditions

$$\mathbf{q}(a) = \mathbf{q}_a,\ \mathbf{q}(b) = \mathbf{q}_b.$$

Standard Mechanical Systems

For a mechanical system

$$M\frac{d^2\mathbf{q}}{dt^2} = -\frac{\partial U}{\partial \mathbf{q}} \tag{25}$$

with positive definite mass matrix M, the Lagrangian is the difference between kinetic energy T and potential energy U:

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) = T(\dot{\mathbf{q}}) - U(\mathbf{q}) = \frac{1}{2}\dot{\mathbf{q}} \cdot M\dot{\mathbf{q}} - U(\mathbf{q}). \tag{26}$$

2.2 Hamiltonian Forms

Standard (real) Hamiltonian Form

The standard Hamiltonian form comes from the Lagrangian form by defining the momentum vector $\mathbf{p}$ with the *Legendre transformation*

$$\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \tag{27}$$

and the *Hamiltonian*

$$\mathcal{H}(\mathbf{p}, \mathbf{q}) = \mathbf{p} \cdot \dot{\mathbf{q}} - \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}). \tag{28}$$

Hamilton's Equations

The differential equations become

$$\frac{dq_n}{dt} = \frac{\partial \mathcal{H}}{\partial p_n}, \quad \frac{dp_n}{dt} = -\frac{\partial \mathcal{H}}{\partial q_n}, \quad 1 \le n \le N, \tag{29}$$

or in vector form

$$\frac{d\mathbf{q}}{dt} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}}, \quad \frac{d\mathbf{p}}{dt} = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}}. \tag{30}$$

Hamiltonian Form for Mechanical Systems

For the mechanical system above, this gives

$$\mathbf{p} = M\dot{\mathbf{q}}, \tag{31}$$

$$\mathcal{H}(\mathbf{p}, \mathbf{q}) = T + U = \frac{1}{2}\mathbf{p} \cdot (M^{-1}\mathbf{p}) + U(\mathbf{q}), \text{ (total energy)} \tag{32}$$

$$\frac{d\mathbf{q}}{dt} = M^{-1}\mathbf{p}, \tag{33}$$

$$\frac{d\mathbf{p}}{dt} = -\frac{\partial U}{\partial \mathbf{q}}. \tag{34}$$

Complex Hamiltonian Form

Some equations, such as ones related to electro-magnetic fields and quantum mechanics, are most naturally stated in terms of complex-valued dependent variables, and then a complex version of Hamiltonian form is more convenient.

For this, the conjugate variable pairs (q_n, p_n) are replaced by pairs of complex quantities (z_n, z_n^*), grouped into two complex vectors $\mathbf{z}$ and $\mathbf{z}^*$.

Although these will be complex conjugates in all applications, the initial development treats them as independent.

Consider a Hamiltonian $\mathcal{H}(\mathbf{z}, \mathbf{z}^*)$ with the restriction that it be real-valued when the conjugate pairs of variables are in fact complex conjugates. The corresponding differential equations are the complex Hamiltonian form

$$\frac{dz_n}{dt} = i\frac{\partial \mathcal{H}}{\partial z_n^*}, \quad \frac{dz_n^*}{dt} = -i\frac{\partial \mathcal{H}}{\partial z_n}. \tag{35}$$

Such a system can be converted to the previous standard Hamiltonian form with the change of variables

$$q_n = \frac{z_n + z_n^*}{\sqrt{2}}, \quad p_n = \frac{z_n - z_n^*}{i\sqrt{2}}. \tag{36}$$

Complex Hamiltonian Form: Only One Equation is Needed

Though $\mathbf{q}$ and $\mathbf{p}$ are in general complex values, restricting them the real values makes the Hamiltonian real and imposes the expected complex conjugacy relation on $\mathbf{z}$ and $\mathbf{z}^*$. Then the second equation in (35) is redundant, and the complex form of the Hamiltonian system is simply

Canonical Complex Hamiltonian Equations

$$\frac{dz_n}{dt} = i\frac{\partial \mathcal{H}}{\partial z_n^*}, \quad \text{or in vector form} \quad \frac{d\mathbf{z}}{dt} = i\frac{\partial \mathcal{H}}{\partial \mathbf{z}^*}. \tag{37}$$

The Discrete Nonlinear Schrödinger Equation

One important illustrative example of complex Hamiltonian form is the The Discrete Nonlinear Schrödinger Equation (3) with Hamiltonian

$$\mathcal{H} = \sum_n \kappa(z_{n+1} - z_n)(z_{n+1}^* - z_n^*) - G(z_n z_n^*).$$

In the practical case of the conjugate variables being complex conjugates the Hamiltonian is

$$\mathcal{H} = \kappa \sum_n |z_{n+1} - z_n|^2 - G(|z_n|^2).$$

Various end conditions can be imposed on the Hamiltonian to deal with the out-of-bounds indices $n = 0$ and $n = N + 1$ that arise, in particular:

- Homogeneous Dirichlet: $z_0 = z_{N+1} = 0$, or
- Periodic: $z_{n+N} = z_n$ (so $z_0 = z_N$ and $z_{N+1} = z_1$).

One can also consider an infinite chain: $n \in \mathbb{N}$, $z_n \to 0$ as $|n| \to \infty$. This is convenient for constructing continuum limit approximations by PDE's.

1. The equations are not in the standard form for a mechanical system,
2. The equations do not separate as $\mathcal{H} = T(\mathbf{p}) + U(\mathbf{q})$, and
3. The Hamiltonian does not arise from a Lagrangian $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})$ via a Legendre transform $\mathbf{p} = \partial\mathcal{L}/\partial\dot{\mathbf{q}}$.

These facts interfere with the use of some conservative numerical methods that work well for other systems of equations: some, but not all of these problems can be overcome.

A General Hamiltonian Form

The above real and complex canonical forms fit a more general form: with the state variables gathered into a vector $\mathbf{y}$,

$$\frac{d\mathbf{y}}{dt} = \mathcal{J} D_{\mathbf{y}} \mathcal{H}(\mathbf{y}) = \mathcal{J}\frac{\partial \mathcal{H}}{\partial \mathbf{y}}(\mathbf{y}) \tag{38}$$

with $\mathcal{J}$ an anti-symmetric matrix.

In the cases of interest here, $\mathcal{J}$ is a constant matrix, but it can also be a matrix-valued function of the state variables, $\mathcal{J}(\mathbf{y})$, so at a few points, the handling of that situation will also be described.

For the real canonical case above,

$$\mathbf{y} = \begin{bmatrix} \mathbf{q} \\ \mathbf{p} \end{bmatrix} = [q_1 \dots q_K \; p_1 \dots p_K]^T, \; \mathcal{J} = J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}. \tag{39}$$

This matrix J is the canonical *symplectic* matrix, and will reappear below.

For the complex canonical case,

$$\mathbf{y} = \mathbf{z} \qquad \mathcal{J} = iJ = i \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}. \tag{40}$$

2.3 Dispersive PDE's in Hamiltonian Form

Dispersive PDE's in Hamiltonian Form

Many important dispersive partial differential equations describing wave motion can be expressed in a Hamiltonian form, and this is the basis of a useful method form constructing conservative discretizations.

For the above example of the nonlinear Schrödinger equation, consider its Hamiltonian as a function of two independent variables, z and z^*:

$$\mathcal{H}(z, z^*) = \int (\nabla z) \cdot (\nabla z^*) - G(zz^*) \, d\mathbf{x}.$$

NLS in Hamiltonian Form

Then denoting the Fréchet derivative with respect to the second variable $\mathbf{z}^*$ as $D_{z^*}\mathcal{H}$, one gets $D_{z^*}\mathcal{H} = -\Delta z - G'(zz^*)z^*$, so the nonlinear Schrödinger equation is

$$\frac{\partial z}{\partial t} = iD_{z^*}\mathcal{H}.$$

One method for conservative spatial discretization of dispersive PDEs to be described in these lectures is based on discretizing such Hamiltonian integrals.

2.4 Conserved Quantities or Invariants, and Noether's Theorem

One starting point for ensuring that *invariants* (also called *conserved quantities* or *first integrals*) of the ODE system are also invariant for a time discretization is a discrete version of Noether's Theorem.

First, some definitions.

Continuous Symmetries and Generators

Definition 1 (Continuous Symmetry on the State Variables)**.** A *continuous symmetry of a Lagrangian* $\mathcal{L}$ is a differentiable one-parameter family of transformations $\mathbf{q} \to g_s(\mathbf{q})$ under which the Lagrangian is invariant:

$$\mathcal{L}(g_s(\mathbf{q}), g_s'(\mathbf{q})\dot{\mathbf{q}}) = \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) \quad \text{for all } s \text{ and all } (\mathbf{q}, \dot{\mathbf{q}}). \tag{41}$$

Note: more general symmetries are possible, but these are the ones relevant to time-discretization methods.

Definition 2 (Generator of a Continuous Symmetry)**.** The *generator* of continuous symmetry g_s is

$$a(\mathbf{q}) := \left.\frac{d}{ds}\right|_{s=0} g_s(\mathbf{q}).$$

Though more general forms are possible, addressing more general symmetries, but for the purposes of numerical methods, we use:

Theorem 3 (A Version of Noether's Theorem)**.** *Consider a system of equations with Lagrangian and Hamiltonian form as above, having a continuous symmetry on the state variables with generator* $a(\mathbf{q})$*. Then the quantity*

$$I(\mathbf{p}, \mathbf{q}) := \mathbf{p} \cdot a(\mathbf{q}) = \left(\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}\right) \cdot a(\mathbf{q}) \tag{42}$$

is an invariant *for solutions of the Euler-Lagrange equations* (24)*.*

That is, for any solution $\mathbf{q}(t)$ *of the Euler-Lagrange equations,*

$$\frac{d}{dt} I(\mathbf{p}(t), \mathbf{q}(t)) = 0.$$

Notes on Noether's Theorem

In practice, the invariants given by this form of Noether's theorem are typically either:

- quadratic in the state variables q_i and p_i, coming from linear transformations (for example, angular momentum which comes from rotational invariance), or
- linear combinations of the state variables, coming from translation groups (for example, linear momentum which comes from translation invariance).

In particular, these are the only symmetries and invariants possible for a mechanical system (25), (26), within the class of continuous symmetries defined above.

It will be seen that this limitation to quadratic invariants is even stronger for time-discretization methods.

The one important conservation law not addressed by Noether's Theorem is conservation of energy, meaning the Hamiltonian. (The exception is for linear systems, with quadratic Hamiltonian.)

On the other hand, this is easy to verify for any system of the general Hamiltonian form of Equation (38), as follows:

$$\begin{aligned}
\frac{d\mathcal{H}}{dt} &= D_{\mathbf{y}}\mathcal{H} \cdot \frac{d\mathbf{y}}{dt} && \text{multivariable chain rule} \\
&= D_{\mathbf{y}}\mathcal{H}(\mathbf{y}) \cdot \mathcal{J} D_{\mathbf{y}}\mathcal{H}(\mathbf{y}) && \text{Hamilton's equations (38)} \\
&= 0 && \text{from the anti-symmetry of } \mathcal{J}.
\end{aligned}$$

2.5 Conservation of Other Invariants in the Examples

In addition to the conserved energy, the above examples all have some other conserved quantities, or *invariants*.

I review the calculations use to verify conservation of these invariants, because these calculations will be mimicked later to show conservation of the same quantities for discrete gradient time-stepping methods.

In many any interesting mechanical systems, the Hamiltonian are invariant under a common translation of all position coordinates, $q_n \to q_n + s$. Such a Hamiltonian can be written in terms of the translation invariant linear combinations $\Delta_{n,m} := q_n - q_m$. This is seen in the molecular chain models above, where the variables q_n appear only in the translation invariant *bond-stretchings* $\Delta_n := q_{n+3} - q_n$.

Combined with the evolution equation $dp_n/dt = -\partial\mathcal{H}/\partial q_n$, this gives conservation of the (linear) momentum $\mathcal{P} = \sum_n p_n$. First,

$$\begin{aligned}\frac{d\mathcal{P}}{dt} = \sum_n \frac{dp_n}{dt} = -\sum_n \frac{d\mathcal{H}}{dq_n} &= -\sum_{n<m}\left[\frac{d\mathcal{H}}{d\Delta_{n,m}}\frac{d\Delta_{n,m}}{dq_n} + \frac{d\mathcal{H}}{d\Delta_{m,n}}\frac{d\Delta_{m,n}}{dq_n}\right] \\ &= -\sum_{n<m}\left[\frac{d\mathcal{H}}{d\Delta_{n,m}}(+1) + \frac{d\mathcal{H}}{d\Delta_{m,n}}(-1)\right].\end{aligned}$$

For any pair of indices a and b with $a \le b$, the term $\dfrac{d\mathcal{H}}{d\Delta_{a,b}}$ occurs twice: once in the first "$(+1)$" term; once in the second "(-1)" term. Thus all terms pair off and cancel out, giving $\dfrac{d\mathcal{P}}{dt} = 0$.

Note: This is the simplest illustration of the famous relationship between continuous symmetries and conservation laws stated in Noether's Theorem, which will be discussed more below.

Phase-shift Invariance and Conservation of the Charge

The DNLS equations and the Davydov exciton-particle model above conserve the *charge*,

$$\mathcal{E} = \sum_n z_n z_n^*, \tag{43}$$

corresponding to the probability density of Schrödinger's equation.

Phase Shift Invariance and its Invariant Quadratic Forms

This is associated to invariance of the Hamiltonian under the linear symmetry

$$z_n \to e^{is} z_n, \quad z_n^* \to e^{-is} z_n^*,$$

and one way to directly verify this conservation law is to note first that the state variables z_n and z_n^* appear in the Hamiltonian only through the quadratic combinations

$$\pi_{n,m} = z_n z_m^*$$

which are invariant under the phase shift. These serve the same purpose as the invariant linear forms $\Delta_{n,m}$ seen above.

Invariance of Charge

To verify invariance of the charge in time, first differentiate:

$$\begin{aligned}
\frac{d\mathcal{E}}{dt} &= \sum_n \frac{d(z_n z_n^*)}{dt} \\
&= \sum_n \frac{dz_n}{dt} z_n^* + \frac{dz_n^*}{dt} z_n \\
&= i \sum_n \frac{\partial \mathcal{H}}{\partial n_k^*} z_n^* - \frac{\partial \mathcal{H}}{\partial z_n} z_n \\
&= i \sum_n \Big[\sum_m \frac{\partial \mathcal{H}}{\partial \pi_{m,n}} \frac{\partial \pi_{m,n}}{\partial z_n^*} z_n^* - \sum_m \frac{\partial \mathcal{H}}{\partial \pi_{n,m}} \frac{\partial \pi_{n,m}}{\partial z_n} z_n \Big] \\
&= i \sum_n \sum_m \Big[\frac{\partial \mathcal{H}}{\partial \pi_{m,n}} z_n z_n^* - \frac{\partial \mathcal{H}}{\partial \pi_{n,m}} z_n^* z_m \Big].
\end{aligned}$$

For any pair of indices a, b with $a \leq b$

- the term $\dfrac{\partial \mathcal{H}}{\partial \pi_{n,m}} z_n z_m^*$ for $m = a$, $n = b$ is $\dfrac{\partial \mathcal{H}}{\partial \pi_{b,a}} z_b z_a^*$, and
- the term $-\dfrac{\partial \mathcal{H}}{\partial \pi_{m,n}} z_n^* z_m$ for $m = b$, $n = a$ is $-\dfrac{\partial \mathcal{H}}{\partial \pi_{b,a}} z_a^* z_b$,

so again all terms pair off and cancel, giving $\dfrac{d\mathcal{E}}{dt} = 0$.

3 Spatial Discretization of PDE's and Continuum Limits

3.1 Discretizations of the Nonlinear Schrödinger Equation

Connecting DNLS and NLS

The discrete nonlinear Schrödinger equation (3) can be connected to the nonlinear Schrödinger equation (2) via a continuum limit in one direction, spatial discretization in the other.

Continuum Limit from DNLS to NLS

The continuum limit applies to the infinite lattice model with n taking all integer values, with the definitions

$$\Delta x := 1/\sqrt{\kappa}, \quad x_n = n\Delta x, \quad z(x_n, t) \approx z_n(t)$$

$$\mathcal{H} = \sum_n \frac{(z_{n+1} - z_n)}{\delta x} \frac{(z_{n+1^*} - z_n^*)}{\delta x} - G(z_n z_n^*)$$

and the limit $\Delta x \to 0$.

It is important for our purposes that the rescaled limit

$$\mathcal{H}[z, z^*] = \lim_{\delta x \to 0} \mathcal{H}(\mathbf{z}, \mathbf{z}^*)\delta x,$$

gives the Hamiltonian functional

$$\mathcal{H}[z, z^*] = \int \frac{\partial z}{\partial x}\frac{\partial z^*}{\partial x} - G(zz^*)\, dx. \tag{44}$$

With this Hamiltonian the nonlinear Schrödinger equation has the form

$$\frac{\partial z}{\partial t} = \mathrm{i}\frac{\delta \mathcal{H}}{\delta z^*}, \tag{45}$$

where the term at right is now a Fréchet derivative.

Notes

- The factor δx in the limit adjusts for the difference that the definition of the Fréchet derivative is in terms of integrals rather than the sums involved in the definition of the gradient. This is convenient, allowing simple conversion between sums and integrals.
- As before, one must formally consider the Hamiltonian as a functional $\mathcal{H}[z, z^*]$ of two independent quantities z and z^*.
- Confirmation of this limit is needed when proceeding in the opposite direction: using numerical solutions of the spatially discrete system as a numerical solution of the original PDE.

Discretizing the Nonlinear Schrödinger Equation

The Hamiltonian form of the Nonlinear Schrödinger Equation allows us to go in the other direction of spatial discretization:

1. Define $x_n := n\delta x$ and $z_n(t) \approx z(x_n, t)$.
2. Approximate the spatial derivative in the Hamiltonian functional (44) with the basic finite difference form
$$\frac{\partial z}{\partial x}(x_n) \approx \frac{z_{n+1} - z_n}{\delta x}.$$
A factor of δx is removed for conversion from Fréchet derivatives to gradients, as explained above.
3. Use this to approximate the Hamiltonian functional $\mathcal{H}[z, z^*]$ by a sum $\mathcal{H}(\mathbf{z}, \mathbf{z})$.
4. Use the resulting Hamiltonian system of ODE's as an approximation of the PDE.

A similar procedure works for higher spatial dimensions as in (2) and for other PDE's with a suitable Hamiltonian form, such as the nonlinear Klein-Gordon equation (1).

Some Conserved Quantities are Preserved, But Not All

The above discretization automatically respects the conservation of (discretized) energy, and also the discretization $\mathcal{E} = \sum_n |z_n|^2$ of the charge $\mathcal{E} = \int |z(x)|^2 \, dx$.

However the discrete system has no counterpart of the momentum

$$\mathcal{P} = \mathrm{i} \int z \frac{\partial z^*}{\partial x} - z^* \frac{\partial z}{\partial x} \, dx.$$

Non-Conservation of Momentum and Traveling Waves

This non-conservation of momentum is manifested in solution of the discrete nonlinear Schrödinger equation by the lack of traveling wave solutions and a slowing of pulse solutions. For example in the the focusing cubic case $G(E) = E^2$, $g(E) = 2E$ of the discrete nonlinear Schrödinger equation there are no solitary wave solutions corresponding to the soliton solutions

$$z(t, x) = A \operatorname{sech}(A(x - vt)) \exp\left[-\frac{\mathrm{i}}{2}(vx - (v^2/4 - A^2)t)\right].$$

of the *one dimensional focusing cubic Schrödinger equation*

$$\mathrm{i}\frac{\partial z}{\partial t} + \frac{\partial^2 z}{\partial x^2} + 2|z|^2 z = 0. \tag{46}$$

Extension to Other Equations

This method extends in an obvious way to other equations such as the nonlinear Schrödinger equation in more space dimensions (2) and the nonlinear Klein-Gordon equation (1).

However, the same inability to conserve both momentum and energy occurs in all cases.

An Alternative Spatial Discretization: The Ablowitz-Ladik System

The underlying problem is that the momentum is related to the continuous symmetry of spatial translation, which no longer exits in the spatial discretization.

There are other spatial discretizations that do conserve discrete counterparts of all three of the above conserved integrals. Most famous is the Ablowitz-Ladik system [**?**], also known as the *integrable discrete nonlinear Schrödinger equation:*

$$\mathrm{i}\frac{dz_n}{dt} + \frac{(z_{n+1} - 2z_n + z_{n-1})}{\delta x^2} \pm |z_n|^2 \frac{z_{n-1} + z_{n-1}}{2} = 0, \tag{47}$$

for the cubic Schrödinger equation with either sign on the cubic nonlinearity (focusing or defocusing). Conservation of momentum requires this be considered as an infinite chain, so with indices $n \in \mathbb{Z}$.

More Momentum Conserving Spatial Discretizations

[Kevrekidis-et-al:2007] introduces a systematic approach to constructing spatial discretizations for the nonlinear Schrödinger equation (2) that conserve the discretized momentum

$$\mathcal{P} = \mathrm{i}\sum_n z_n(z^*_{n+1} - z^*_{n-1}) - z^*_n(z_{n+1} - z_{n-1})$$

as well as discretizations of $\mathcal{H}$ and $\mathcal{E}$.

For the one dimensonal focusing cubic case above, $G(E) = E^2$, this method gives the family of discretizations

$$\mathrm{i}\frac{dz_n}{dt} + \frac{z_{n+1} - 2z_n + z_{n-1}}{\delta x^2} + \left[(1-\alpha)|z_n|^2 + \frac{\alpha}{2}(|z_{n-1}|^2 + |z_{n+1}|^2)\right]\frac{z_{n-1} + z_{n+1}}{2} = 0.$$

This includes the Ablowitz-Ladik system as $\alpha = 0$.

However, the resulting ODE systems are not in canonical Hamiltonian form.

For example, the Ablowitz-Ladik system has Hamiltonian form with a non-constant symplectic matrix $\mathcal{J}(\mathbf{z})$ and a quite different Hamiltonian. There is a nonlinear transformation of variables that puts it in real canonical Hamiltonian form but still with a different Hamiltonian. However as we will see soon, the original energy (the standard Hamiltonian) is not conserved by any of the time discretization methods discussed in these lectures, because it is not quadratic.

3.2 Spatial Discretization of the Hamiltonian and Higher Order Methods

The method of spatial discretization of the Hamiltonian be applied to a variety of dispersive wave equations with a suitable canonical Hamiltonian form. Also, methods with higher order of accuracy are easily constructed.

The simple forward difference approximation of $\partial z/\partial x$ used above leads to ODEs with the basic three point discretization of the second derivative

$$\frac{\partial^2 z}{\partial x^2} \approx \frac{z_{n-1} - 2z_n + z_{n+1}}{(\Delta x)^2}.$$

More accurate spatial discretizations of derivatives in the Hamiltonian can be used.

Centered Differences and a Fourth Order Accurate ODE Approximation

For example, using the centered difference approximation

$$\frac{\partial z}{\partial x}(x_n) \approx \frac{z_{n+1} - z_{n-1}}{2\Delta x}$$

gives an ODE system with fourth order accurate five-point difference approximation of $\partial^2 z/\partial x^2$.

More General Lattice Equations

Combining this with the possibility of higher spatial dimension in the PDE leads to a more general class of discrete nonlinear Schrödinger equations given by Hamiltonians

$$\mathcal{H} = \sum_n \left[\sum_m l_{nm} z_n z_m^* + G(z_n z_n^*) \right]$$

for symmetric matrix $L = \{l_{nm}\}$.

The ODEs are then a general family of *lattice nonlinear Schrödinger equations*

$$\mathrm{i}\frac{dz_n}{dt} + \sum_m l_{nm} z_m + g(z_n z_n^*) z_n = 0. \tag{48}$$

This form includes, for example, the two-dimensional lattice nonlinear Schrödinger equation (7)

3.3 Space-Time Discretization of Conservation Law Forms

Several authors ([Strauss+Vázquez:1978, Jiménez:1994, **?**] have developed a discrete difference calculus approach for directly constructing energy conserving space and time discretizations of the nonlinear Klein-Gordon equation (1).

This has been extended in [Jiménez:1994] to the nonlinear Schrödinger equation (2) and a related family of nonlinear Dirac equations. All the PDEs have a simple "single point" nonlinearity coming from a potential G added to the Hamiltonian of the corresponding linear equation.

The method uses a space and time discretization of the conservation law form

$$\frac{\partial u}{\partial t} = v, \quad \frac{\partial}{\partial t}\left[\frac{1}{2}v^2 + \frac{1}{2}\left|\nabla u\right|^2 - G(u)\right] = \nabla \cdot (v \nabla u)\,.$$

The resulting methods conserve a spatial discretization of the energy, and in the case of the nonlinear Schrödinger equation also conserves the discretized charge $\mathcal{E}$. (For the nonlinear Dirac equation, its two charges are conserved. Here the name "charge" indicates a conserved quantity for a PDE that is the integral of a quadratic function of the unknowns, not involving any derivatives: these are the easiest to discretize in both space and time!)

However, none of these methods solves the problem already seen above: none also conserves the momentum.

For the nonlinear Klein-Gordon equation, it is possible to start instead from the conservation law form for momentum:

$$\frac{\partial u}{\partial t} = v, \quad \frac{\partial}{\partial t}\left(v \nabla u\right) = \frac{\partial}{\partial t}\left[\frac{1}{2}v^2 + \frac{1}{2}\left(\nabla u\right)^2 - G(u)\right].$$

This discretization then conserves momentum, but not energy.

Although the methods proceed directly from the PDE to a fully discrete system, they give results that can also be attained with a version of the two stage process seen in these lectures:

1. Discretize in space through discretizing the Hamiltonian (using the most basic method above, giving only second order accuracy in space),
2. Discretize the resulting ODE system in time using an energy-momentum conserving method (using one of the basic methods described below, again giving only second order accuracy in time.)

This two step approach adds options with increased order of accuracy in both space and time.

Features of the Resulting ODE Systems

The systems of ODE's $\frac{d\mathbf{y}}{dt} = \mathbf{f}(\mathbf{y})$ arising from such discretizations of PDEs typically have the following features:

1. The number of unknowns, N, is very large.
2. If the PDE is semi-linear or quasi-linear, the ODE system inherits an analogous property like the form
$$\frac{d\mathbf{y}}{dt} = \mathbf{f}(\mathbf{y}) = L(\mathbf{y}) + \mathbf{g}(\mathbf{y})$$
where L is a constant matrix and the nonlinear term $\mathbf{g}(\mathbf{y})$ is in some sense relatively small. For example, Jacobian matrix $D_{\mathbf{y}}\mathbf{g}(\mathbf{y})$ of far smaller norm than L for relevant values of $\mathbf{y}$.
3. The coupling between unknowns is *local*, or *predominantly local*, where
   - *local* means that the $N \times N$ Jacobian matrix $A = D_{\mathbf{y}}\mathbf{f}(\mathbf{y})$ is *banded*: the only non-zero elements A_{ij} are ones with $|i - j| \leq w$ for some *bandwidth* $w \ll N$: w is far smaller than N.
   - *predominantly local* means that the system is quasi-linear as above and the linear part L has low bandwidth w.

For example, the above spatial discretization of the nonlinear Schrödinger is quasi-linear because $\kappa = 1/(\Delta x)^2 \gg 1$ makes the linear term dominant, and it is local with bandwidth $w = 1$.

4 Variational Time-Stepping Methods, Symmetric Runge-Kutta Methods, and Conservation of Quadratic Invariants

4.1 Variational Time-Stepping Methods: Discretizing Hamilton's Principle

The connection between continuous symmetries and conserved quantities stated in Noether's Theorem has a natural counterpart in time discretizations for systems that have a Lagrangian form, through discretization of Hamilton's Principle.

To discretize, first divide the time interval $a \leq t \leq b$ with nodes

$$t_\tau = a + \tau h,\ h = \frac{(b-a)}{T},\ 0 \leq \tau \leq T,$$

and discretize the state variable with $\mathbf{q}_\tau \approx \mathbf{q}(t_\tau)$. Then the action integral is approximated by a sum

$$\mathcal{S}_h\left(\left\{\mathbf{q}_\tau\right\}_{\tau=0}^{T}\right) = \sum_{\tau=1}^{T} \mathcal{L}_h(\mathbf{q}_{\tau-1}, \mathbf{q}_\tau)$$

using approximations of the action integral on each time step

$$\mathcal{L}_h(\mathbf{q}_{\tau-1}, \mathbf{q}_\tau) \approx \int_{t_{\tau-1}}^{t_\tau} \mathcal{L}(\mathbf{q}(t), \dot{\mathbf{q}}(t))\, dt. \tag{49}$$

The discretized equations to solve are the *discrete Euler-Lagrange equations* for a minimum (or critical point) of the discrete action:

$$D_2\mathcal{L}_h(\mathbf{q}_{\tau-1}, \mathbf{q}_\tau) + D_1\mathcal{L}_h(\mathbf{q}_\tau, \mathbf{q}_{\tau+1}) = 0. \tag{50}$$

To solve these, introduce the discrete momenta

$$p_\tau := -D_1\mathcal{L}_h(\mathbf{q}_\tau, \mathbf{q}_{\tau+1}).$$

Combining this with (50) and increasing all indices by 1 leads to the variational method as pair of equations

$$\mathbf{p}_\tau = -D_1\mathcal{L}_h(\mathbf{q}_\tau, \mathbf{q}_{\tau+1}), \quad \mathbf{p}_{\tau+1} = D_2\mathcal{L}_h(\mathbf{q}_\tau, \mathbf{q}_{\tau+1}). \tag{51}$$

Thus, starting with values $\mathbf{q}_\tau, \mathbf{p}_\tau$, the time stepping procedure is to solve the first of these for $\mathbf{q}_{\tau+1}$, and then use the second to evaluate $\mathbf{p}_{\tau+1}$. It can be shown that this has a unique solution for small enough values of h, and so then defines a map $(\mathbf{q}_\tau, \mathbf{p}_\tau) \to (\mathbf{q}_{\tau+1}, \mathbf{p}_{\tau+1})$.

Some More Notation

Before continuing, some further notation will be useful.

We will focus on the time advance map for single time step, from a time t to $t+h$. Thus for a scalar variable x, a vector $\mathbf{x}$, and likewise for other variables like $\mathbf{q}$, $\mathbf{p}$, and $\mathbf{z}$:

- h or δt denotes the change in t over the time step.
- x alone without arguments denotes the value $x(t)$ at time t, typically the beginning of the current time step.
- x^+ denotes the value $x(t+h)$.
- $\delta x = x^+ - x$.
- X or $\overline{x} = \dfrac{x+x^+}{2}$.
- $\dot{X} = \dfrac{\delta x}{\delta t}$.

The Variational Method Time-Stepping Map

With this notation, the variational method in (51) becomes the equations

$$\mathbf{p} = -D_1\mathcal{L}_h(\mathbf{q},\mathbf{q}^+), \quad \mathbf{p}^+ = D_2\mathcal{L}_h(\mathbf{q},\mathbf{q}^+) \tag{52}$$

determining the time-stepping map $(\mathbf{q},\mathbf{p}) \to (\mathbf{q}^+,\mathbf{p}^+)$.

4.2 A Discrete Noether's Theorem

Noether's Theorem has a discrete analog. (This and most other theorems on numerical methods can be found in the book [GNI] by E. Hairer, C. Lubich and G. Wanner, herein referred to as "GNI".)

Theorem 4 (A Discrete Noether's Theorem (GNI Theorem VI.6.7))**.** *Suppose the one parameter group of transformations $\mathbf{q} \to g_s(\mathbf{q})$ also leaves the discrete Lagrangian invariant:*

$$\mathcal{L}_h(g_s(\mathbf{q}), g_s(\mathbf{q}^+)) = \mathcal{L}_h(\mathbf{q},\mathbf{q}^+) \quad \textit{for all } s \textit{ and all pairs } (\mathbf{q},\mathbf{q}^+). \tag{53}$$

Then the discrete system (52)

$$\mathbf{p} = -D_1\mathcal{L}_h(\mathbf{q},\mathbf{q}^+), \quad \mathbf{p}^+ = D_2\mathcal{L}_h(\mathbf{q},\mathbf{q}^+)$$

again has the invariant $I(\mathbf{p},\mathbf{q}) := \mathbf{p}\cdot a(\mathbf{q})$.
That is, $(\mathbf{p}^+)\cdot a(\mathbf{q}^+) = \mathbf{p}\cdot a(\mathbf{q})$.

Notes on the Discrete Noether Theorem

1. Not all invariants of the ODE system are invariants of the discrete system, because invariance of the Lagrangian as in (41) does not ensure invariance of the discrete Lagrangian as in (53). In fact, in general this is at best true for linear symmetry groups, and so only for quadratic invariants. For example, no Runge-Kutta method can in general conserve polynomial invariants of order greater than two.
2. Some systems like DNLS and the Davydov models do not have a Lagrangian form, so we need to see how to express the discrete variational equations in terms of the Hamiltonian alone.

4.3 The Implicit Midpoint, Störmer-Verlet Methods, Other Runge-Kutta Methods

The Variational Method Using Midpoint Quadrature

As a first example of the procedure of constructing a variational method and converting to purely Hamiltonian form, we consider one of the simplest yet most important methods, and its conversion to Hamiltonian form for the special case of a standard mechanical system.

We start with a discrete Lagrangian given by combining the midpoint rule with the most basic derivative approximation $\frac{d\mathbf{q}}{dt} \approx \dot{\mathbf{Q}} = \frac{\mathbf{q}^+ - \mathbf{q}}{h}$ along with $\mathbf{Q} = \frac{\mathbf{q} + \mathbf{q}^+}{2}$:

$$\mathcal{L}_h(\mathbf{q}, \mathbf{q}^+) = h\mathcal{L}\left(\mathbf{Q}, \dot{\mathbf{Q}}\right) \tag{54}$$

The pair of equations to solve, first for $\mathbf{q}^+$ and then for $\mathbf{p}^+$, are

$$\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}\left(\mathbf{Q}, \dot{\mathbf{Q}}\right) - \frac{h}{2}\frac{\partial \mathcal{L}}{\partial \mathbf{q}}\left(\mathbf{Q}, \dot{\mathbf{Q}}\right) \tag{55}$$

$$\mathbf{p}^+ = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}\left(\mathbf{Q}, \dot{\mathbf{Q}}\right) + \frac{h}{2}\frac{\partial \mathcal{L}}{\partial \mathbf{q}}\left(\mathbf{Q}, \dot{\mathbf{Q}}\right) \tag{56}$$

The Midpoint Variational Method for Mechanical Systems

For the standard mechanical system of (26) and (32)

this becomes

$$M\frac{\mathbf{q}^+ - \mathbf{q}}{h} = \mathbf{p} - \frac{h}{2}D_{\mathbf{q}}U\left(\frac{\mathbf{q} + \mathbf{q}^+}{2}\right) \tag{57}$$

$$\mathbf{p}^+ = M\frac{\mathbf{q}^+ - \mathbf{q}}{h} - \frac{h}{2}D_{\mathbf{q}}U\left(\frac{\mathbf{q} + \mathbf{q}^+}{2}\right) \tag{58}$$

The Implicit Midpoint Method

This is just the second order accurate *implicit midpoint method*

$$\frac{\mathbf{y}^+ - \mathbf{y}}{h} = f\left(\frac{\mathbf{y} + \mathbf{y}^+}{2}\right) \tag{59}$$

applied to the Hamiltonian system

$$\frac{d}{dt}\begin{bmatrix} \mathbf{q} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} D_{\mathbf{p}}\mathcal{H} \\ -D_{\mathbf{q}}\mathcal{H} \end{bmatrix} = \begin{bmatrix} M^{-1}\mathbf{p} \\ -D_{\mathbf{q}}U(\mathbf{q}) \end{bmatrix}.$$

The time-stepping method also has the suggestive form

$$\dot{\mathbf{Q}} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}}(\mathbf{Q}, \mathbf{P}), \tag{60}$$

$$\dot{\mathbf{P}} = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}}(\mathbf{Q}, \mathbf{P}). \tag{61}$$

Conservation of Quadratic Invariants

For a linear symmetry group $\mathbf{q} \to A(s)\mathbf{q}$, given by a matrix-valued one-parameter group $A(s)$, invariance of the Lagrangian gives invariance of the above discrete Lagrangian (54)

$$\mathcal{L}_h(\mathbf{q}, \mathbf{q}^+) = h\mathcal{L}\left(\frac{\mathbf{q}+\mathbf{q}^+}{2}, \frac{\mathbf{q}^+ - \mathbf{q}}{h}\right)$$

so that for example the components of angular momenta, which are the invariants given by invariance of the Lagrangian under rotations about each coordinate axis, are also invariant under this discrete scheme.

Time Symmetry, Symplecticity, and Partitioned Runge-Kutta Methods

To construct conservative time-stepping methods for Hamiltonian systems that do not come from a Lagrangian form and yet conserve all quadratic (and linear) invariants, three concept are useful:

- Symmetric methods (under time-reversal)
- symplectic maps and methods, and
- partitioned Runge-Kutta methods.

Definition 5 (Symmetric Time-Stepping Method)**.** A time stepping method is *symmetric* if the time-stepping map of step size $-h$ is the inverse of the map for step size h.

For example, the implicit midpoint method is symmetric — but no explicit one-step method is.

From now on, we will require that all methods are symmetric, because this is empirically seen to have desirable effects on the long-term behavior of numerical solutions, and is also convenient when constructing methods of higher order accuracy.

Symplectic Maps and Time-Stepping Methods

Definition 6 (Symplectic Linear Mapping)**.** A linear mapping $A : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is *symplectic* if

$$A^T J A = J.$$

with J the canonical symplectic matrix $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$ introduced in (39).

Definition 7 (Symplectic Mapping)**.** A differentiable map $g : U \to \mathbb{R}^{2n}$, U an open subset of $\mathbb{R}^{2n}$, is *symplectic* if its Jacobian is symplectic at each point of U.

Definition 8 (Symplectic Time-Discretizations)**.** A time-discretization method for Hamiltonian systems is *symplectic* if the time advance map $(\mathbf{q}, \mathbf{p}) \to (\mathbf{q}^+, \mathbf{p}^+)$ is symplectic.

Theorem 9. *A mapping $(\mathbf{q}, \mathbf{p}) \to (\mathbf{q}^+, \mathbf{p}^+)$ is symplectic if and only if there is a sufficiently smooth* generating function $S(\mathbf{q}, \mathbf{q}^+)$ *for which*

$$\mathbf{p} = -\frac{\partial S}{\partial \mathbf{q}}(\mathbf{q}, \mathbf{q}^+), \quad \mathbf{p}^+ = \frac{\partial S}{\partial \mathbf{q}^+}(\mathbf{q}, \mathbf{q}^+).$$

Any variational method (51)

$$\mathbf{p} = -D_1\mathcal{L}_h(\mathbf{q}, \mathbf{q}^+), \quad \mathbf{p}^+ = \partial_2\mathcal{L}_h(\mathbf{q}, \mathbf{q}^+)$$

is of this form with generating function $S = \mathcal{L}_h$, so

Theorem 10. *Every variational method is symplectic, (and conversely).*

General Runge-Kutta Methods

All variational methods of practical interest are based on *Runge-Kutta methods*.

For a general autonomous system $\frac{d\mathbf{q}}{dt} = \mathbf{f}(\mathbf{q})$, a general s-stage Runge-Kutta method is of the form

$$\dot{\mathbf{Q}}_i = \mathbf{f}(\mathbf{Q}_i, \mathbf{P}_i), \qquad 1 \le i \le s, \tag{62}$$

$$\mathbf{Q}_i = \mathbf{q} + h \sum_{j=1}^{s} a_{ij} \dot{\mathbf{Q}}_\tau, \qquad 1 \le i \le s, \tag{63}$$

$$\mathbf{q}^+ = \mathbf{q} + h \sum_{i=1}^{s} b_i \dot{\mathbf{Q}}_i. \tag{64}$$

with all b_i non-zero and $\sum_{i=1}^{s} b_i = 1$. For example, the implicit midpoint method is a very simple case with $s = 1$, $b_1 = 1$, and $a_{11} = 1/2$.

Symplectic Runge-Kutta Methods and Quadratic Invariants

For Runge-Kutta methods, there is simple relationship between conserving quadratic (and linear) invariants and being symplectic, and also an algebraic criterion to test for these properties.

Theorem 11 (GNI Theorems IV.2.2, VI.4.3, VI.7.10)**.** *For a Runge-Kutta method applied to a Hamiltonian system, the following are equivalent:*

1. *The coefficients satisfy*
$$b_i a_{ij} + b_j a_{ji} = b_i b_j, \quad 1 \le i, j \le s. \tag{65}$$
2. *The method is symplectic.*
3. *The method conserves any quadratic invariant $I(\mathbf{q}, \mathbf{p}) = \mathbf{p} \cdot A\mathbf{q}$ of the Hamiltonian system.*

Also, all linear invariants of a Hamiltonian system are conserved by any Runge-Kutta method.

Explicit Runge-Kutta Methods Cannot Be Symplectic

An *explicit* Runge-Kutta method is one in which the $\dot{\mathbf{Q}}_i$ and $\mathbf{Q}_i$ are given successively by simple evaluation, without equation solving: this is true if and only if $a_{ij} = 0$ whenever $i \leq j$.

It is then impossible to satisfy the above criterion for being symplectic. Thus only implicit Runge-Kutta methods are of interest to us.

However, there is sometimes a way around this, as we see next.

For the canonical Hamiltonian equations with their partitioned form

$$\frac{d\mathbf{q}}{dt} = \mathbf{f}(\mathbf{q},\mathbf{p}) = \frac{\partial \mathcal{H}}{\partial \mathbf{p}}, \quad \frac{d\mathbf{p}}{dt} = \mathbf{g}(\mathbf{q},\mathbf{p}) = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}},$$

it is natural to consider the more general class of

Partitioned Runge-Kutta Methods

$$\dot{\mathbf{Q}}_i = \mathbf{f}(\mathbf{Q}_i, \mathbf{P}_i), \qquad \dot{\mathbf{P}}_i = \mathbf{g}(\mathbf{Q}_i, \mathbf{P}_i). \tag{66}$$

$$\mathbf{Q}_i = \mathbf{q} + h\sum_{j=1}^{s} a_{ij}\dot{\mathbf{Q}}_i, \qquad \mathbf{P}_i = \mathbf{p} + h\sum_{j=1}^{s} \hat{a}_{ij}\dot{\mathbf{P}}_i, \tag{67}$$

$$\mathbf{q}^+ = \mathbf{q} + h\sum_{i=1}^{s} b_i\dot{\mathbf{Q}}_i, \qquad \mathbf{p}^+ = \mathbf{p} + h\sum_{i=1}^{s} \hat{b}_i\dot{\mathbf{P}}_i. \tag{68}$$

There is an equivalence much as above:

Theorem 12 (GNI Theorems IV.4.6 and IV.2.4)**.** *For a partitioned Runge-Kutta method applied to a Hamiltonian system, the following are equivalent:*

1. *The coefficients satisfy*

$$b_i\hat{a}_{ij} + \hat{b}_j a_{ji} = b_i\hat{b}_j, \qquad 1 \leq i,j \leq s, \tag{69}$$

$$b_i = \hat{b}_i, \qquad 1 \leq i \leq s. \tag{70}$$

2. *The method is symplectic.*

3. *The method conserves any quadratic invariant $I(\mathbf{q},\mathbf{p}) = \mathbf{p}\cdot A\mathbf{q}$ of the Hamiltonian system.*

Also, the condition (70) *alone guarantees that all linear invariants are conserved.*

From now on, we will always assume the condition $\hat{b}_i = b_i$. Thus, we have little need to discuss linear invariants further.

Contrary to the situations with standard (non-partitioned) Runge-Kutta methods, *there are explicit symplectic partitioned Runge-Kutta methods.* We will see an important example soon.

Variational Methods of Partitioned Runge-Kutta Form

To construct partitioned Runge-Kutta methods for Hamiltonian systems, we restrict our attention to discrete Lagrangians that are of the s-stage Runge-Kutta form

$$\mathcal{L}_h(\mathbf{q},\mathbf{q}^+) = h\sum_{i=1}^{s} b_i\mathcal{L}(\mathbf{Q}_i,\dot{\mathbf{Q}}_i) \tag{71}$$

with all b_i non-zero, sum $\sum_i b_i = 1$, and

$$\mathbf{Q}_i = \mathbf{q} + h \sum_{j=1}^{s} a_{ij} \dot{\mathbf{Q}}_i. \tag{72}$$

For a variational method, the $\dot{\mathbf{Q}}_i$ extremize this discrete Lagrangian (71) subject to the constraint

$$\mathbf{q}^+ = \mathbf{q} + h \sum_{i=1}^{s} b_i \dot{\mathbf{Q}}_i. \tag{73}$$

It is clear that applying any linear transformation group to $\mathbf{q}$ leads to $\mathbf{Q}_i$ and $\dot{\mathbf{Q}}_i$ that are transformed in the same way, and thus when the Lagrangian is invariant under a linear symmetry, so is any such discrete Lagrangian. Hence the Discrete Noether's Theorem ensures:

Theorem 13. *A discrete variational method of Runge-Kutta form conserves all quadratic and linear invariants of the corresponding Lagrangian ODE system.*

Symplectic Partitioned Runge-Kutta Methods for Hamiltonian Systems

For a discrete variational method in Runge-Kutta form as above, the three "q" equations on the left-hand side of the partitioned form are already satisfied due to equations (72) and (73).

Thus, defining

$$\dot{\mathbf{P}}_i = \frac{\partial \mathcal{L}}{\partial \mathbf{q}}(\mathbf{Q}_i, \dot{\mathbf{Q}}_i), \quad \mathbf{P}_i = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}(\mathbf{Q}_i, \dot{\mathbf{Q}}_i)$$

and imposing the conditions (69), (70),

$$b_i \hat{a}_{ij} + \hat{b}_j a_{ji}, = b_i \hat{b}_j, \quad b_i = \hat{b}_i, 1 \le i \le s$$

one gets a symplectic partitioned Runge-Kutta method.

Thus for *any* Runge-Kutta method, choosing

$$\hat{b}_i = b_i, \quad \hat{a}_{ij} = b_j(1 - a_{ji}/b_i), \tag{74}$$

gives a symplectic method that conserves quadratic invariants:

$$\dot{\mathbf{Q}}_i = \frac{\partial \mathcal{H}}{\partial \mathbf{p}}(\mathbf{Q}_i, \mathbf{P}_i), \qquad \dot{\mathbf{P}}_i = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}}(\mathbf{Q}_i, \mathbf{P}_i),$$

$$\mathbf{Q}_i = \mathbf{q} + h \sum_{j=1}^{s} a_{ij} \dot{\mathbf{Q}}_i, \qquad \mathbf{P}_i = \mathbf{p} + h \sum_{j=1}^{s} \hat{a}_{ij} \dot{\mathbf{P}}_i,$$

$$\mathbf{q}^+ = \mathbf{q} + h \sum_{i=1}^{s} b_i \dot{\mathbf{Q}}_i, \qquad \mathbf{p}^+ = \mathbf{p} + h \sum_{i=1}^{s} b_i \dot{\mathbf{P}}_i.$$

The Implicit Midpoint Method is a Symplectic Method for General Hamiltonian Systems

The implicit midpoint method above gives a particularly simple and elegant case, because the only new coefficient is $\hat{a}_{11} = 1/2 = a_{11}$.

With this and the condition $\hat{b}_i = b_i$ already imposed, this new method is not really partitioned, but is just a normal Runge-Kutta method: the implicit midpoint method again.

Thus the implicit midpoint method is now shown to be a symplectic method for Hamiltonian systems in general, not just for the case of mechanical systems seen above.

Since symplectic form is guaranteed by the above algebraic conditions, this also extends to the more general Hamiltonian form $d\mathbf{y}/dt = \mathcal{J} D_{\mathbf{y}} \mathcal{H}$ of (38).

The Variational Method with Trapezoid Rule Quadrature

One other basic variational/symplectic method is given by using the trapezoid rule in the discrete Lagrangian in place of the midpoint rule:

$$\mathcal{L}_h(\mathbf{q}, \mathbf{q}^+) = \frac{h}{2}\mathcal{L}\left(\mathbf{q}, \frac{\mathbf{q}^+ - \mathbf{q}}{h}\right) + \frac{h}{2}\mathcal{L}\left(\mathbf{q}^+, \frac{\mathbf{q}^+ - \mathbf{q}}{h}\right)$$

This gives

$$\begin{aligned}
\mathbf{p} &= \frac{1}{2}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}\left(\mathbf{q}, \frac{\mathbf{q}^+ - \mathbf{q}}{h}\right) + \frac{1}{2}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}\left(\mathbf{q}^+, \frac{\mathbf{q}^+ - \mathbf{q}}{h}\right) - \frac{h}{2}\frac{\partial \mathcal{L}}{\partial \mathbf{q}}\left(\mathbf{q}, \frac{\mathbf{q}^+ - \mathbf{q}}{h}\right) \\
\mathbf{p}^+ &= \frac{1}{2}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}\left(\mathbf{q}, \frac{\mathbf{q}^+ - \mathbf{q}}{h}\right) + \frac{1}{2}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}\left(\mathbf{q}^+, \frac{\mathbf{q}^+ - \mathbf{q}}{h}\right) + \frac{h}{2}\frac{\partial \mathcal{L}}{\partial \mathbf{q}}\left(\mathbf{q}^+, \frac{\mathbf{q}^+ - \mathbf{q}}{h}\right)
\end{aligned}$$

Hamiltonian form for the Trapezoid Rule Variational Method

The Hamiltonian version of this can be expressed as

$$\begin{aligned}
\mathbf{P} &= \mathbf{p} - \frac{h}{2}\frac{\partial \mathcal{H}}{\partial \mathbf{q}}(\mathbf{P}, \mathbf{q}), \\
\mathbf{q}^+ &= \mathbf{q} + \frac{h}{2}\left(\frac{\partial \mathcal{H}}{\partial \mathbf{p}}(\mathbf{P}, \mathbf{q}) + \frac{\partial \mathcal{H}}{\partial \mathbf{p}}(\mathbf{P}, \mathbf{q}^+)\right), \\
\mathbf{p}^+ &= \mathbf{p} - \frac{h}{2}\frac{\partial \mathcal{H}}{\partial \mathbf{q}}(\mathbf{P}, \mathbf{q}^+),
\end{aligned} \tag{75}$$

or an alternative version with the roles of $\mathbf{q}$ and $\mathbf{p}$ reversed.

The Störmer-Verlet Method

This method is often called the *Störmer-Verlet method*: like the midpoint method, this is second order accurate, symplectic, and time-reversal symmetric.

Historically the Störmer-Verlet method was introduced only for the case of a standard mechanical system by C. Störmer in 1903 for astronomical calculations, and independently by L. Verlet in 1967 for molecular dynamics.

However, the first publication of the method was actually by Newton, in the *Principia*!

Mechanical Systems: the Leap-Frog Method

In that case of mechanical systems, the Störmer-Verlet method can be rewritten as the explicit 2-step *leap-frog method*

$$M\frac{\mathbf{q}_{n+1} - 2\mathbf{q}_n + \mathbf{q}_{n-1}}{h^2} = -D_{\mathbf{q}}U(\mathbf{q}_n).$$

More generally, for a system with separable Hamiltonian $\mathcal{H} = T(\mathbf{p}) + U(\mathbf{q})$, this method is explicit:

$$\begin{aligned}
\mathbf{P} &= \mathbf{p} - \frac{h}{2}\frac{\partial U}{\partial \mathbf{q}}(\mathbf{q}) \\
\mathbf{q}^+ &= \mathbf{q} + h\frac{\partial T}{\partial \mathbf{p}}(\mathbf{P}) \\
\mathbf{p}^+ &= \mathbf{p} - \frac{h}{2}\frac{\partial U}{\partial \mathbf{q}}(\mathbf{q}^+)
\end{aligned}$$

Notes on the Störmer-Verlet/Leap-Frog Method

Since conservative and symplectic methods are typically implicit, including all symplectic Runge-Kutta methods, this explicit symplectic partitioned Runge-Kutta method is very attractive. It is thus popular for use with classical mechanics, such as in celestial mechanics and molecular biology.

On the other hand, it does not handle stiff systems well, whereas the implicit midpoint method is unconditionally stable (A-stable and B-stable).

Also, its advantages are lost with non-separable Hamiltonians, such as for discrete nonlinear Schrödinger equations.

4.4 Higher Order Accuracy by Symmetric Step Composition

The methods seen so far are only second order accurate in time.

Fortunately, the method of *symmetric step composition* [Creutz+Gocksch:1989, Forest:1989, Suzuki:1990, Yoshida:1990] gives a systematic way to construct methods of any higher even order while preserving all the interesting properties: conservation of quadratic invariants, time-reversal symmetry, and symplectic form.

The idea is to start with a basic symmetric time-stepping method (in particular, the mid-point or Störmer-Verlet method) and construct a composite step from a sequence of s basic steps, of lengths

$$\beta_1 h, \ldots, \beta_s h.$$

The obvious consistency condition is that

$$\beta_1 + \cdots + \beta_s = 1.$$

If in addition the basic method is of even order p then choosing the *step size factors* β_i with

$$\beta_1^{p+1} + \cdots + \beta_s^{p+1} = 0$$

and the symmetry condition

$$\beta_{s+1-i} = \beta_i, \quad 1 \le i \le s$$

ensures that the composite step is also time-reversal symmetric, and of even order $p + 2$.

The process can then be repeated, constructing composites of composite steps to achieve arbitrarily high order.

The Triple-Jump and Suzuki Fractal Methods

The simplest such method is the *triple-jump*, with three steps:

$$\beta_1 = \beta_3 = \frac{1}{2 - 2^{1/p+1}}, \quad \beta_2 = 1 - 2\beta_1 = \frac{2^{1/p+1}}{2 - 2^{1/p+1}}.$$

Unfortunately, the steps go out of the range from t_τ to $t_{\tau+1}$, as all step size factors are greater than one in magnitude.

For example, with $p = 2$ as for the midpoint method, $\beta_1 \approx 1.35$ and $\beta_2 \approx -1.70$.

Thus a better approach is the five-step *Suzuki fractal method,* where all $|\beta_i| < 1$:

$$\beta_1 = \beta_2 = \beta_4 = \beta_5 = \frac{1}{4 - 4^{1/p+1}}, \quad \beta_3 = 1 - 4\beta_1 = \frac{4^{1/p+1}}{4 - 4^{1/p+1}}.$$

For $p = 2$, the factors are $\beta_1 \approx 0.41$, $\beta_2 \approx -0.66$.

Diagonally Implicit Runge-Kutta [DIRK] Methods

Amongst symmetric (and thus implicit) Runge-Kutta methods, the equation solving needed is greatly simplified if one can solve sequentially for $\mathbf{Q}_1$, then $\mathbf{Q}_2$, and so on, so that one solves s successive smaller systems of simultaneous equations.

This is the case for *Diagonally Implicit Runge-Kutta [DIRK] methods*: ones with

$$a_{ij} = 0 \quad \text{whenever } i < j. \tag{76}$$

Such methods are closely related to step composition methods and the implicit midpoint method:

Theorem 14 (GNI, Theorem VI.4.4)**.** *Any symplectic DIRK method is given by the symmetric composition of midpoint steps with time steps given by* $\beta_i = b_i$.

Thus in particular any symplectic DIRK method is symmetric.

4.5 Higher Order Accuracy with Gaussian Integration Methods

The previous results leave only the following possibilities for time-stepping methods that conserve quadratic invariants:

1. Symmetric compositions of implicit midpoint rule steps.
2. Partitioned Runge-Kutta methods, such as the Störmer-Verlet method or compositions based on it. (These however are problematic for non-separable Hamiltonian systems which do not partition naturally, such as DNLS or the Davydov system.)
3. Runge-Kutta methods that are not diagonally implicit.
4. Methods that are not at all of Runge-Kutta form. (This will be seen to be the only possibility that allows exact conservation of energy.)

We now consider the last two of these possibilities.

Gaussian Methods

Amongst non-diagonally implicit Runge-Kutta methods, the most important examples are the Gaussian methods.

These are a sequence of s-stage, symmetric, symplectic A-stable and B-stable implicit Runge-Kutta methods of order $2s$, built on the corresponding s-point Gauss-Legendre quadrature method.

For $s = 1$, one gets the implicit midpoint method yet again.

For $s = 2$ one gets the fourth-order accurate method with

$$b_1 = b_2 = \frac{1}{2} \tag{77}$$

$$a_{11} = a_{22} = \frac{1}{4}, \quad a_{12} = \frac{1}{4} + \frac{\sqrt{3}}{6}, \quad a_{12} = \frac{1}{4} - \frac{\sqrt{3}}{6}. \tag{78}$$

The latter is often recommended as the best general purpose choice of time-stepping method for Hamiltonian systems, except when one needs very high order accuracy.

4.6 Runge-Kutta Methods Cannot Conserve Both Energy and Quadratic Invariants

Runge-Kutta Methods Cannot Conserve Both Energy and Quadratic Invariants

There is an unfortunate limitation on all Runge-Kutta methods, and all the above variational and symplectic methods: *no such method can in general conserve both quadratic invariants and the energy (Hamiltonian).*

The main reason is a theorem of [Ge+Marsden:1988] which essentially says that a symplectic method cannot also conserve energy. More precisely, a symplectic method that also conserves energy is either the exact time stepping map except for shifts in time, or a system that decomposes into several parts of this form.

The Exception: Conservation of Energy for Linear Systems

Energy will be conserved for a linear system of equations, because the energy itself is then a quadratic invariant. This case is rather trivial, and can effectively be reduced to the system of uncoupled single component systems $\mathcal{H} = \sum_n \left(\frac{p_n^2}{2m_n} + \gamma_n q_n^2 \right)$.

The energy is simply the sum of the quadratic invariants $I_n = \frac{p_n^2}{2m_n} + \gamma_n q_n^2$, which are the energies for the uncoupled sub-systems.

Any symplectic method will conserve the energy for each sub-system (q_n, p_n), meaning that each sub-solution moves along the correct elliptical or hyperbolic path, every point of which is on the trajectory of the true solution of the ODE for that pair. Thus at worst, the numerical solution for each (q_n, p_n) at time t_τ errs by corresponding to the true solution at a slightly different time.

As seen above, any Runge-Kutta method (including the more general class of partitioned Runge-Kutta methods) that conserves quadratic momenta is symplectic, and so had this inability to conserve energy.

Thus we turn next to a different approach, based on energy conservation and using the general Hamiltonian form $d\mathbf{y}/dt = \mathcal{J} D_{\mathbf{y}} \mathcal{H}$, not relying on any partitioned or separable form or on the Lagrangian form.

5 Energy-Momentum Conserving Methods

Just as the derivation of variational methods is based on approximating integrals in Hamilton's Principle, an elegant approach to deriving methods that exactly conserve energy (the Hamiltonian) is to approximate the gradients appearing in the Hamiltonian equation

$$\frac{d\mathbf{y}}{dt} = \mathcal{J} D_{\mathbf{y}} \mathcal{H}.$$

Again,

- conserving other invariants requires an appropriate choice of discretization, and
- there is a natural limitation to quadratic (and linear) invariants.

5.1 Energy-Momentum Methods for Central Force Mechanical Systems

[**?**] introduced time discretization methods that conserve energy, angular, and linear momentum for N particle central force mechanical systems with Hamiltonian

$$\mathcal{H} = \frac{1}{2} \sum_n \frac{\|\mathbf{p}_n\|^2}{2m_n} + \sum_{m>n} U_{n,m}(\|\mathbf{q}_n - \mathbf{q}_m\|) \tag{79}$$

where now each $\mathbf{q}_n(t)$ and $\mathbf{p}_n(t)$ is a vector in $\mathbb{R}^3$. The same methods were developed in a different way and then extended to continuum mechanics by Juan Simo et al, in particular in [Simo-et-al:1992].

However, the extensions to PDE's are essentially equivalent to first applying a suitable spatial discretization as seen above, giving a system of ODE's in the above form, and then using a time discretization equivalent to the discrete gradient method to be seen next. Thus, details will be omitted here.

5.2 Discrete Gradient Methods

Discrete Gradient Methods

The basic idea, originating in the work of O. Gonzalez and Juan Simo [Gonzales:1996, Gonzales+Simo:1996] is to define a discrete approximation of the gradient

$$D_{\mathbf{y}} f(\mathbf{y}) \approx \tilde{D}_{\mathbf{y}} f(\mathbf{y}, \mathbf{y}^+) = \left\langle \tilde{D}_1 f(\mathbf{y}, \mathbf{y}^+), \tilde{D}_2 f(\mathbf{y}, \mathbf{y}^+), \dots \right\rangle, \tag{80}$$

giving the *discrete Hamilton's equation*

$$\frac{\delta \mathbf{y}}{\delta t} = \mathcal{J} \tilde{D}_{\mathbf{y}} \tilde{\mathcal{H}}(\mathbf{y}, \mathbf{y}^+). \tag{81}$$

In the case of a variable matrix $\mathcal{J}(\mathbf{y})$, one instead uses an approximation $\tilde{\mathcal{J}}(\mathbf{y}, \mathbf{y}^+)$. The natural choice is the midpoint approximation $\tilde{\mathcal{J}}(\mathbf{y}, \mathbf{y}^+) = \mathcal{J}(\overline{\mathbf{y}})$.

The approach presented here follows [LeMesurier:2012a, LeMesurier:2012b].

From now on we will assume the natural consistency condition

$$\lim_{\mathbf{y}^+ \to \mathbf{y}} \tilde{D}_{\mathbf{y}} f(\mathbf{y}, \mathbf{y}^+) = D_{\mathbf{y}} f(\mathbf{y}).$$

A discrete gradient cannot simply be constructed from independently defined discrete approximations of the partial derivatives, because an important relation must be imposed on the components: all discrete gradients are required to satisfy the

Discrete Multivariable Chain Rule

$$\delta f = \tilde{D}_{\mathbf{y}} f(\mathbf{y}, \mathbf{y}^+) \cdot \delta \mathbf{y}, = \sum_n \tilde{D}_n f(\mathbf{y}, \mathbf{y}^+) \delta y_n. \tag{82}$$

A suitable discrete Jacobian $\tilde{D}_{\mathbf{y}} \mathbf{f}(\mathbf{y}, \mathbf{y}^+)$ for a vector-valued function $\mathbf{f}$ is then given by combining discrete gradients of each component function of $\mathbf{f}$.

Conservation of Energy

Conservation of energy is easily shown, by mimicking the argument used above for a (continuous time) Hamiltonian system:

$$\begin{aligned}
\frac{\delta \mathcal{H}}{\delta t} &= \tilde{D}_{\mathbf{y}} \mathcal{H}(\mathbf{y}, \mathbf{y}^+) \cdot \frac{\delta \mathbf{y}}{\delta t} && \text{discrete multivariable chain rule} \\
&= \tilde{D}_{\mathbf{y}} \mathcal{H}(\mathbf{y}, \mathbf{y}^+) \cdot \mathcal{J} \tilde{D}_{\mathbf{y}} \mathcal{H}(\mathbf{y}, \mathbf{y}^+) && \text{discrete Hamilton's equations (81)} \\
&= 0 && \text{from the anti-symmetry of } \mathcal{J}.
\end{aligned}$$

Choosing a Discrete Gradient

There are infinitely many choices of discrete gradient satisfying this chain rule condition (loosely: if you choose all but one component, the chain rule determines the value of that last component).

The challenge, as with variational methods, is to choose one that gives *conservation of invariants*: the new strategy here is to choose a discrete gradient that allows verification of conservation by *mimicking the calculations that work for differential equations.*

Note that this is what was just done for conservation of energy.

There seems to be no universal way to do this, with the problem in particular being that

associativity of products cannot be respected

so that triple products cannot be handled in a canonical way.

On the other hand, there is a canonical approach if one avoids triple products, and this leads again to the familiar result that

quadratic and linear invariants can be handled,

but not in general invariants of other forms.

For a wide range of systems arising from physical problems, this is enough.

Difference Calculus for Functions of One Variable

For functions $f(x)$ of a single variable, the discrete multivariable chain rule (82) dictates a simple and familiar difference scheme:

$$\tilde{D}_x f(x, x^+) := \begin{cases} \dfrac{\delta f}{\delta x}, & x^+ \neq x \\ \dfrac{df}{dx}(x), & x^+ = x. \end{cases} \tag{83}$$

That is, the standard centered difference approximation, with the exact derivative used when needed.

This approximation is at best second order accurate, and leads to the basic discrete gradient method being second order accurate, as for the implicit midpoint method. However, step composition methods can again be used to construct higher order methods.

Avoiding Division By Zero

A greater problem is that this will be used with x^+ unknown, so it is unknown which of the two forms above applies.

Thus it is highly desirable (but not quite essential) to simplify the first form so as to eliminate division by δx, to get a universally valid formula.

To start with, this is possible for natural number powers with

$$\tilde{D}_x(x^r)(x, x^+) = \frac{\delta(x^r)}{\delta x} = (x)^{r-1} + (x)^{r-2}(x^+) + \cdots + (x^+)^{r-1}.$$

Linear Combinations, Compositions, and Inverses

Sums and constant factors are naturally handled by linearity, and there is a canonical choice for a *discrete chain rule* for compositions $(f \circ g)(\mathbf{x})$ with f a function of one variable:

$$\tilde{D}_n(f \circ g)(\mathbf{x}, \mathbf{x}^+) = \tilde{D}_g f(g, g^+) \tilde{D}_n g(\mathbf{x}, \mathbf{x}^+) \tag{84}$$

or

$$\tilde{D}_{\mathbf{y}}(f \circ g)(\mathbf{x}, \mathbf{x}^+) = \tilde{D}_g f(g, g^+) \cdot \tilde{D}_{\mathbf{x}} g(\mathbf{x}, \mathbf{x}^+). \tag{85}$$

From this, the inverse of a function of one variable, $y = f^{(-1)}(x)$, is handled as:

$$\tilde{D}_x y = \tilde{D}_x \left(f^{(-1)} \right)(x, x^+) = \frac{1}{\tilde{D}_y f(y, y^+)}.$$

Combined with the above result for $f(x) = x^r$ and linearity, all elementary rational functions of one variable can be handled with a single formula, without the special "division by zero" case.

Products: A Little Thought Required, at Last

For a product of two factors, the possibilities for a product rule satisfying the discrete chain rule condition include

$$\delta(fg) = \alpha(g\delta f + f^+\delta g) + (1 - \alpha)(g^+\delta f + f\delta g)$$

but the most reasonable choice is $\alpha = 1/2$, which is the only choice that has time-reversal symmetry and respects commutativity of products. It also has the advantage of giving second order accuracy.

The Time-Symmetric Discrete Product Rule

Thus we use adopt the *Discrete Product Rule* based on decomposing $\delta(fg)$ as

$$\delta(fg) = \overline{g}\delta f + \overline{f}\delta g. \tag{86}$$

That is, for a product $f(\mathbf{x})g(\mathbf{x})$,

$$\tilde{D}_n(fg)(\mathbf{x},\mathbf{x}^+) = \overline{g}(\tilde{D}_n f)(\mathbf{x},\mathbf{x}^+) + \overline{f}(\tilde{D}_n g)(\mathbf{x},\mathbf{x}^+), \tag{87}$$

$$\tilde{D}_{\mathbf{x}}(fg)(\mathbf{x},\mathbf{x}^+) = \overline{g}(\tilde{D}_{\mathbf{x}} f)(\mathbf{x},\mathbf{x}^+) + \overline{f}(\tilde{D}_{\mathbf{x}} g)(\mathbf{x},\mathbf{x}^+). \tag{88}$$

For Linear Systems, this is the Implicit Midpoint Rule, and so is Unconditionally Stable

For linear systems, the Hamiltonian is quadratic and so this leads to the implicit midpoint method.

Thus, any choice of discrete gradient method following these rules will have the unconditional stability of the midpoint method.

Non-Associativity of the Product Rule

For products of more than two factors, it is impossible to construct a generally applicable rule. If one applies the rule above to product uvw via the various factorizations $u(vw)$, $(uv)w$, $v(uw)$ etc., one gets different results, with different consequences for conservation of invariants.

Also, symmetrizations such as averaging over all alternatives do not necessarily give a form that respects the symmetries and conservation laws of the Hamiltonian.

This is a Calculus for Formulas, not Functions

Instead, the approach here is limited to handling conservation laws whose verification only depends on the handling of terms involving at most products of two variables — specifically, the invariant quadratic (or linear) combinations of state variables noted above.

This means that:

the difference calculus defined here applies to formulas,
with the order of evaluation of all operations specified.

Theorem 15. *If a Hamiltonian system* (38)

(a) has a collection of conserved quadratic invariants (depending only on state variables), and

(b) the associated symmetry group of the Hamiltonian consists of affine transformations of the variables, and

(c) the Hamiltonian is "manifestly invariant" in that it can be expressed entirely in terms of quadratic combinations $\hat{y}_a(\mathbf{y})$ of the original variables that are invariant under this symmetry group

$$\mathcal{H}(\mathbf{y}) = \hat{\mathcal{H}}(\hat{\mathbf{y}}(\mathbf{y})) \tag{89}$$

then we get a time discretization that conserves these invariants by using the discrete gradient given by applying the discrete chain rule (85) *to formula* (89).

A Conservative Discrete Gradient Scheme

The resulting time-stepping method is

$$\frac{\mathbf{y}^+ - \mathbf{y}}{\delta t} = \mathcal{J}\tilde{D}_{\hat{\mathbf{y}}}\hat{\mathcal{H}}(\hat{\mathbf{y}}, \hat{\mathbf{y}}^+) \cdot \tilde{D}_{\mathbf{y}}\hat{\mathbf{y}}(\mathbf{y}, \mathbf{y}^+). \tag{90}$$

Note:

- The choice of the discrete gradient $\tilde{D}_{\hat{\mathbf{y}}}\hat{\mathcal{H}}(\hat{\mathbf{y}}, \hat{\mathbf{y}}^+)$ in the first factor does not matter.
- The second factor $\tilde{D}_{\mathbf{y}}\hat{\mathbf{y}}(\mathbf{y}, \mathbf{y}^+)$ is determined by the discrete product rule (87) and linearity, through terms like

  $$\tilde{D}_{y_a}(y_a y_b) = \overline{y}_b \quad (a \neq b), \qquad \tilde{D}_{y_a}((y_a)^2) = 2\overline{y}_a$$

  Thus the formula for $\tilde{D}_{\mathbf{y}}\hat{\mathbf{y}}(\mathbf{y}, \mathbf{y}^+)$ comes from the formula for the exact gradient $\tilde{D}_{\mathbf{y}}\hat{\mathbf{y}}(\mathbf{y})$ through the substitution $\mathbf{y} \to \overline{\mathbf{y}}$.

Proof Idea

The main idea in the proof of this result is that for any conserved quantity $Q(\mathbf{y})$ of the ODE system that is quadratic in the state variables y_j,

$$\frac{dQ}{dt} = D_{\mathbf{y}}Q(\mathbf{y}) \cdot \mathcal{J}D_{\hat{\mathbf{y}}}\hat{\mathcal{H}} \cdot D_{\mathbf{y}}\hat{\mathbf{y}} = 0 \tag{91}$$

and this fact does not depend on the details of the function $\hat{\mathcal{H}}$.

Why? Because any choice of $\hat{\mathcal{H}}$ gives a Hamiltonian $\mathcal{H}(\mathbf{y})$ with the required symmetry properties and thus makes Q a conserved quantity, so that the above equation is true. The result depends only on the terms $D_{\mathbf{y}}Q(\mathbf{y})$ and $D_{\mathbf{y}}\hat{\mathbf{y}}$, which are linear in the y_j.

With $\delta Q/\delta t$ for the discrete gradient scheme, the only change in these linear gradient terms is the substitution $y_j \to \overline{y}_j$, which is merely a "renaming", and does not affect the validity of the identity: $\delta Q/\delta t = 0$, so $Q(\mathbf{y})$ is conserved by the time-discrete system.

Example: the LaBudde-Greenspan Method

This method can be applied to the central force particle system of Hamiltonian (79)

by rewriting each potential term as $U_{n,m} = W_{n,m}(\|q_n - q_m\|^2)$. Because then, the Hamiltonian is expressed entirely in terms of the quadratic combinations
$\|q_n - \mathbf{q}_m\|^2 = (q_{n,1} - \mathbf{q}_{m,1})^2 + (q_{n,2} - \mathbf{q}_{m,2})^2 + (q_{n,3} - \mathbf{q}_{m,3})^2$,
$\|\mathbf{p}_n\|^2 = (\mathbf{p}_{n,1})^2 + (\mathbf{p}_{n,2})^2 + (\mathbf{p}_{n,3})^2$
which are invariant under the rotational and translation symmetries associated with angular and linear momentum.

The resulting method is that of [**?**] and [Simo-et-al:1992].

5.3 Practical Implementation: an Iterative Solution Method

A Linearly Implicit Iterative Scheme for Solving Discrete Gradient Systems

The system of equations (90) will be nonlinear (unless the Hamiltonian system itself is linear), so we need an iterative solution method.

The following method preserves the linear stability properties and exact momentum conservation:

- Set $\mathbf{y}^{(0)} = \mathbf{y}$.
- Construct successive approximations $\mathbf{y}^{(k)}$ of $\mathbf{y}^+$ by solving

$$\mathbf{y}^{(k+1)} - \mathbf{y} = h\mathcal{J}\tilde{D}_{\hat{\mathbf{y}}}\hat{\mathcal{H}}(\hat{\mathbf{y}}, \hat{\mathbf{y}}^{(k)}) \cdot \tilde{D}_{\mathbf{y}}\hat{\mathbf{y}}(\mathbf{y}, \mathbf{y}^{(k+1)}), \tag{92}$$

  a variant of (90).

That is, the nonlinear part $D_{\hat{\mathbf{y}}}\hat{\mathcal{H}}$ is approximated with the current best available approximation $\mathbf{y}^{(k)}$ of $\mathbf{y}^+$, while the linear terms are left in terms of the unknown $\mathbf{y}^{(k+1)}$ to be solved for.

Exact Conservation of Quadratic Invariants at Each Iteration

This equation is linear in the unknown $\mathbf{y}^{(k+1)}$, making its solution straightforward, and much as above, we have:

Theorem 16. *Each iterate* $\mathbf{y}^{(k)}$ *given by the above scheme* (92) *conserves all quadratic first integrals that are conserved by the original discrete gradient scheme* (90)*.*

The key is again the irrelevance of the form of the approximation of $\tilde{D}_{\hat{\mathbf{y}}}\hat{\mathcal{H}}$, which is the only term that changes between the original scheme and this iterative scheme.

Unconditional Linear Stability

Another advantage of this approach to iterative solution is that it has unconditional linear stability, since for a linear system, $\tilde{D}_{\hat{\mathbf{y}}}\hat{\mathcal{H}}$ is constant, the scheme converges in a single iteration, and is the unconditionally stable implicit midpoint method.

This second order accuracy after one iteration for linear systems is seen in practice to carry over in part to mildly nonlinear systems: for various test cases, the accuracy at each iteration is about what one would expect with one more iteration of a standard explicit predictor-corrector approach.

Energy Conservation: Only in the Limit $k \to \infty$

Energy however is only conserved in the limit as the iterates $\mathbf{y}^{(k)}$ converge to $\mathbf{y}^+$.

However, assuming this convergence, sufficient iterations will give energy accuracy far greater than for a scheme that does not conserve energy, such as a symplectic scheme.

Iterating until energy is accurate within rounding error is typically practical: if this take too many iterations, it is better for overall accuracy to reduce the time step size h to speed the convergence.

Handling General Elementary Function Hamiltonians

When this iterative method is applied for a non-algebraic Hamiltonian, the difficulties noted above with the definition of $D_x f(x, x^+)$ in (83) can be avoided, because the problem arises only in the approximation $\tilde{D}_{\hat{y}}\hat{\mathcal{H}}(\hat{y}, \hat{y}^{(k)})$ at each iteration, and this depends only on the already-known quantities y and $y^{(k)}$. Thus one knows whether to use the difference quotient form or the exact derivative at each iteration.

More carefully, the exact derivative should be used if the relevant denominator is smaller than some threshold related to rounding error.

5.4 Higher order accuracy through symmetric step composition

Symmetric step composition can again be used to construct methods of higher order accuracy in time. Thus, every symplectic DIRK method has an energy-momentum conserving counterpart.

However, there is no known counterpart of the higher order Gaussian symplectic methods.

5.5 The Semi-implicit Methods of Strauss, Vázquez et al

The PDE discretization methods of Strauss, Vázquez et al described previously use time discretizations that rely on the ODE system having the form

$$\frac{d\mathbf{y}}{dt} = L\mathbf{y} + \mathbf{g}(\mathbf{y})$$

with L an anti-Hermitian matrix and point-wise nonlinearity $g_n(\mathbf{y}) = G_n'(y_n)$.

This allows a splitting approach, where the two terms at right are handled differently.

Essentially, one combines a conservative method for each:

1. For the linear part, any symplectic method is also energy conserving, so either the implicit midpoint rule, leap-frog or other options can be used: the original Strauss-Vázquez method for the nonlinear Klein-Gordon equation method uses leap-frog.
2. For the nonlinear part, which is a decoupling system of single variable ODE's, the discrete gradient method applies very simply.

The original Strauss-Vázquez method, generalized to second order ODE systems

$$\frac{d^2 y_n}{dt^2} = \sum_m L_{n,m} y_m + G_n'(y_n),$$

is

$$\frac{y_n^{\tau+1} - 2y_n^{\tau} + y_n^{\tau-1}}{2\delta t} + \sum_m L_{n,m} y_m^{\tau} + \frac{G_n(y_n^{\tau+1}) - G_n(y_n^{\tau-1})}{y_n^{\tau+1} - y_n^{\tau-1}}.$$

This has the advantage of being *semi-implicit*: one gets a separate nonlinear equation to solve for each $y_n^{\tau+1}$ instead of a large, coupled system of nonlinear equations.

On the other hand, its stability properties are not good: variants with better stability properties for large time steps are derived by replacing the linear part $L\mathbf{y}^{\tau}$ by forms like $L\left(\frac{1}{4}\mathbf{y}^{\tau-1} + \frac{1}{2}\mathbf{y}^{\tau} + \frac{1}{4}\mathbf{y}^{\tau+1}\right)$ [Li+VuQuoc:1995].

Using $L\left(\frac{1}{2}\mathbf{y}^{\tau-1} + \frac{1}{2}\mathbf{y}^{\tau+1}\right)$ gives the above A-stable discrete gradient method, but jumping two time steps at a time.

The Strauss-Vázquez Method Adapted to the Nonlinear Schrödinger Equation

[Jiménez:1994] adapts these methods to the general nonlinear Schrödinger equation, first discretized in space as

$$i\frac{dz_n}{dt} + \sum_m L_{n,m} z_m + G'(|z_n|^2) z_n = 0$$

and then discretizing in time to get the semi-implicit method

$$i\frac{z_n^{\tau+1}-z_n^{\tau-1}}{2\delta t}+\sum_m L_{n,m}z_m^{\tau}+\frac{G(|z_n^{\tau+1}|^2)-G(|z_n^{\tau-1}|^2)}{|z_n^{\tau+1}|^2-|z_n^{\tau-1}|^2}\cdot\frac{z_n^{\tau+1}+z_n^{\tau-1}}{2}=0.$$

Notes on the Method of Jiménez

The stiffness of spatial discretizations of the nonlinear Schrödinger equation make this a poor method for solving such ODEs, due to its severe time step size restriction.

However, if one is studying an inherently discrete nonlinear Schrödinger equation model, where solutions are not necessarily slowly varying in space and the highest spatial frequency modes are relevant, this method can be attractive, due to the easier, decoupled form of the nonlinear equations to be solved.

5.6 A Multi-gradient Method for Conserving All Invariants

Multi-gradient System Form

[**?**] introduces an extension of the discrete gradient method in which each member of a collection of conserved quantities is given the same status as the Hamiltonian in a new *multi-gradient system* form, leading to conservation of each by the same simple argument as was used for conservation of energy above.

To help with notation, note that a general Hamiltonian system can be expressed as

$$\frac{dy_n}{dt}=\sum_m \mathcal{J}_{nm}D_m\mathcal{H}$$

where $\mathcal{J}$ is anti-symmetric: $\mathcal{J}_{nm}=-\mathcal{J}_{mn}$.

For an autonomous system of equations

$$\frac{d\mathbf{y}}{dt}=\mathbf{f}(\mathbf{y})$$

with a collection of invariants $I_1(\mathbf{y}), I_2(\mathbf{y}), \ldots, I_r(\mathbf{y})$, the main step is to express the differential equations in terms of all of their gradients as

$$\frac{dy_n}{dt}=\sum_{n_1,\ldots,n_r} A_{n,n_1,n_2,\ldots,n_r}D_{n_1}I_1(\mathbf{y})\ldots D_{n_r}I_r(\mathbf{y}) \tag{93}$$

where now $A=A(\mathbf{y})$ is a function whose value at each $\mathbf{y}$ is a totally anti-symmetric $(r+1)$-tensor: interchanging the values of any two of its $r+1$ indices negates the value.

Note that to conserve energy and momenta for a Hamiltonian system, the Hamiltonian itself is one of these invariants.

This can be done so long as invariants are independent, in that their gradients

$$\mathbf{v}_i=D_{\mathbf{y}}I_i$$

are linearly independent at each point $\mathbf{y}$. To do so, first define the symmetric positive definite matrix B with entries $B_{ij}=\mathbf{v}_i\cdot\mathbf{v}_i$. Then a suitable tensor A is

$$A=\frac{1}{\det B}\mathbf{f}\wedge\mathbf{v}_1\cdots\wedge\mathbf{v}_r.$$

With the ODE system in this form, it is straightforward to verify that all the invariants are conserved by the *discrete multi-gradient method*

$$\frac{\delta y_n}{\delta t} = \sum_{n_1,\ldots,n_r} \overline{\mathcal{J}}_{n,n_1,n_2,\ldots,n_r} \tilde{D}_1 I_1(\mathbf{y},\mathbf{y}^+)\ldots\tilde{D}_r I_r(\mathbf{y},\mathbf{y}^+)$$

with any choice of the discrete gradient and any totally anti-symmetric approximation $\overline{A}(\mathbf{y},\mathbf{y}^+) \approx A(\mathbf{y})$. As before, the natural choice is the midpoint approximation $\overline{\mathcal{J}}(\mathbf{y},\mathbf{y}^+) = \mathcal{J}(\overline{\mathbf{y}})$, giving second order accuracy.

Disadvantages of the Multi-Gradient Method: Severe Nonlinearity and Non-Locality

Although this method has the great advantage of being able to conserve invariants of arbitrary form, this comes at the cost of a highly nonlinear form, due to that determinant in the denominator: for a Hamiltonian system, that determinant depends on the gradient of the Hamiltonian $D_{\mathbf{y}}\mathcal{H}$. The determinant also means that the coupling of unknowns is in general non-local: the linear equations involved in iterative solutions will in general be be full, not banded.

Thus, this approach might be well suited to highly nonlinear systems of ODE's that have non-quadratic invariants other than the Hamiltonian, but for the large quasi-linear systems described above, it loses the quasi-linearity that is so useful in solving the discrete gradient method, and also for the midpoint method.

6 Numerical Tests

Numerical Tests

We conclude with a few basic comparisons of the discrete gradient method to the implicit midpoint method, illustrating some advantages for the discrete gradient method in preserving qualitative features of solutions.

The equations solved are the nonlocal discrete nonlinear Schrödinger equation (6)

As mentioned before, a continuum limit approximation of assuming slow variation along the chain leads to the cubic focusing nonlinear Schrödinger equation

$$i\frac{\partial z}{\partial t} + \frac{\partial^2 z}{\partial x^2} + 2|z|^2 z = 0$$

with solutions including the hyperbolic secant solitons

$$\psi(t,x) = A\,\mathrm{sech}(A(x-vt))\exp\left[-\frac{i}{2}(vx-(v^2/4-A^2)t)\right].$$

so initial data is chosen to see how closely solutions of the DNLS system resemble these solitons.

Specifically, initial data is of the form above except of double the height, so that in NLS a soliton plus other radiation is expected to develop.

However, accuracy is tested without relying on the continuum limit, but comparing to an effectively exact solution computed by using time step so small that further reduction produces no visible change, and the results agree between methods.

Exercises

1. Implement the fourth-order Gauss method, a fourth order symplectic DIRK method, and the two fourth order discrete gradient methods given by the the triple jump and Suzuki fractal methods for this system, and repeat the comparisons made here.

2. Repeat with some of the other sample equations above.

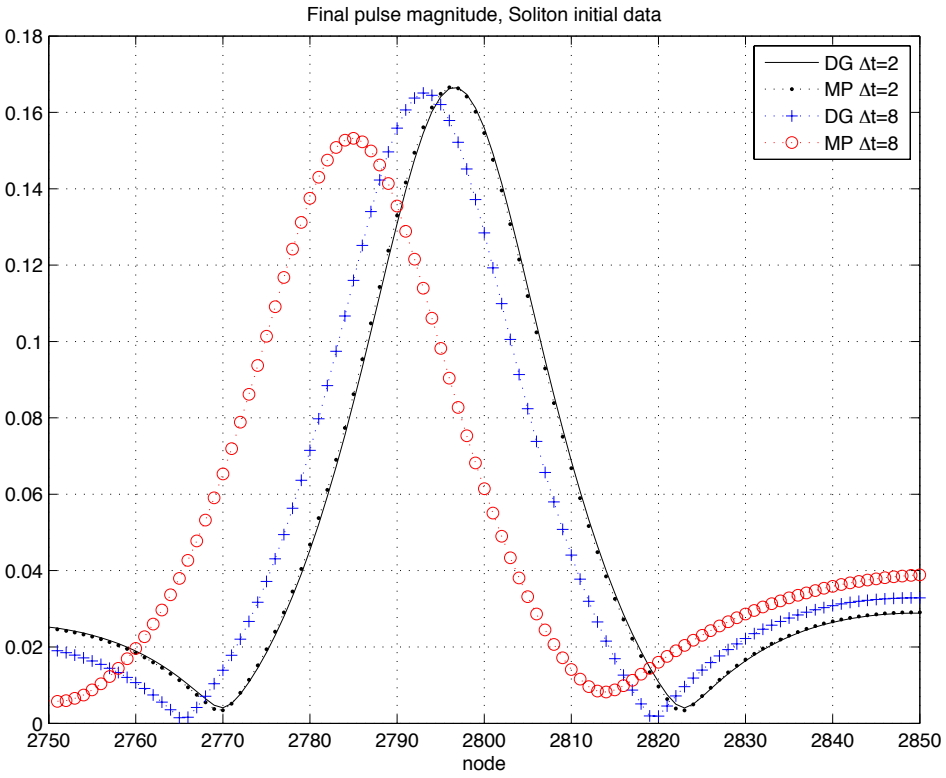

Figure 1: NLDNLS, $A = 2$: $|z_n|$ at time $t = 40000$.

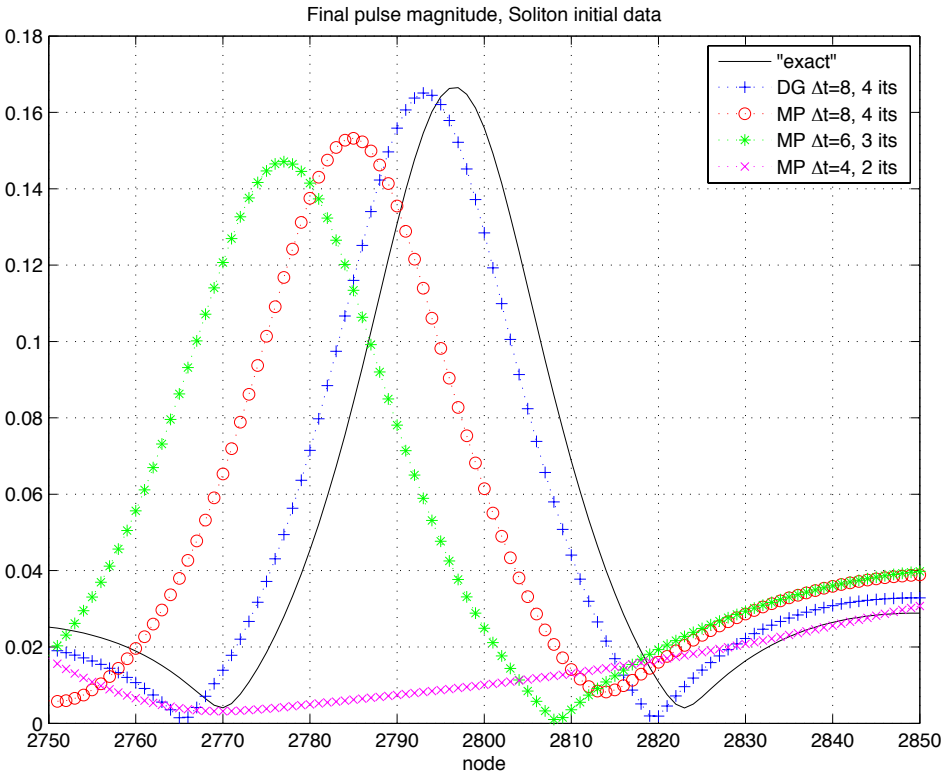

Figure 2: NLDNLS, $A = 2$: $|z_n|$ at time $t = 40000$, methods of roughly equal time cost.

References

[Davydov:1971] A. S. Davydov. *Theory of molecular excitations*. Plenum Press, New York, 1971.

[Davydov+Kislukha:1973] A. S. Davydov and N. I. Kislukha, *Solitary excitations in one-dimensional molecular chains*, Phys. Stat. Sol. B 59 (1973) 465–

[Ablowitz+Ladik:1975] M. J. Ablowitz and J. F. Ladik, *Nonlinear differential-difference equations,* Journal of Mathematical Physics 16 (1975) 598–603.

[LaBudde+Greenspan:1976] R. A. LaBudde and D. Greenspan Numer. Meth. 25 (1976) 323–346 and 26 (1976) 1–26.

[Strauss+Vázquez:1978] W. Strauss and L. Vázquez *Numerical solution of a nonlinear Klein-Gordon equation*, Journal of Computational Physics 28 (2),1978, 271-278.

[Ge+Marsden:1988] Z. Ge and J. E. Marsden Phys. Lett. A 133 (1988) 134–139.

[Creutz+Gocksch:1989] M. Creutz and A. Gocksch, *HIgher order hybrid Monte-Carlo algorithms*, Phys. Rev. Lett. 63 (1989) 9–12.

[Forest:1989] E. Forest *Canonical integrators as tracking codes*, AIP Conference Proceedings 184 (1989) 1106–1136.

[Suzuki:1990] M. Suzuki Phys. Lett. A 146 (1990) 319–323.

[Yoshida:1990] H. Yoshida Phys. Lett. A 150 (1990) 262–268.

[Simo-et-al:1992] Juan C. Simo, N. Tarnow and K. K. Wong Comput. Methods Appl. Mech. Eng. 100 (1992) 63–116.

[Jiménez:1994] S. Jiménez *Derivation of the discrete conservation laws for a family of finite difference schemes*, Appl. Math. and Comp. 64 (1994) 13–45.

[Li+VuQuoc:1995] Shaofan Li and Loc Vu-Quoc, *Finite difference calculus invariant structure of a class of algorithms for the nonlinear Klein-Gordon equation*, SIAM Journal on Numerical Analysis 32(6) (1995) 1839-1875.

[Gonzales:1996] O. Gonzales *Time integration and discrete Hamiltonian systems*, J, Nonlinear Sc. 6 (1996) 449–467.

[Gonzales+Simo:1996] O. Gonzales and Juan C. Simo *On the Stability of Symplectic and Energy-Momentum Algorithms for Nonlinear Hamiltonian Systems with Symmetry*, Comput. Methods Appl. Mech. Eng. 134 (1996) 197–222.

[McLaughlin-et-al:1999] R. I. McLaughlin, G. R. W. Quispel and N. Robidoux *Geometric integration using discrete gradients*, Philos. Trans. R. Soc. Lond. Ser. A, 357 (1999) 1021–1045.

[GNI] D. Hairer, C. Lubich and G. Wanner, *Geometric Numerical Integration*, Second Edition, Springer Velag 2006.

[Kevrekidis-et-al:2007] P. G. Kevrekidis, S. V. Dmitriev and A. A. Sukhorukov, *On a class of spatial discretizations of equations of the nonlinear Schrödinger type*, Mathematics and Computers in Simulation 74 (2007) 343–351.

[LeMesurier:2012a] B. LeMesurier *Conservative, Unconditionally Stable Discretization Methods for Hamiltonian Equations, Applied to Wave Motion in Lattice Equations Modeling Protein Molecules*, Physica D 241 (2012) 1–10.

[LeMesurier:2012b] B. LeMesurier *Studying Davydov's ODE Model of Wave Motion in α-helix Protein Using Exactly Energy-Momentum Conserving Discretizations for Hamiltonian Systems*, Mathematics and Computers in Simulation 82 (2012), 1239-1248.

Further Reading

[1] A. S. Davydov, *Biology and Quantum Mechanics*, Pergamon Press, Oxford, 1982.

[2] Qian-Shun Chang and Lin-Bau Xu, A numerical method for a system of generalized nonlinear Schrödinger equations. *J. Comp. Math.*, 4(3):191–199, 1986. (Note: the earliest energy-momentum method for NLS I know of, but see below.)

[3] P. G. Kevrekidis, *On a class of discretizations of Hamiltonian nonlinear partial differential equations*, Physica D 183 (2003) 68-86.

[4] M. Delfour, M. Fortin and G. Payre, *Finite-difference solutions of a nonlinear Schrödinger equation*, J. Comp. Phys 44 (1981) 277-288. (Note: possibly the earliest energy-momentum method for NLS.)

[5] J. M. Sanz-Serna, *Methods for the numerical solution of the nonlinear Schrödinger equation*, Math. Comp. (1985) 21-27. (Note: cites and generalizes method of D+F+P:1981; handles power law NLS)