Conservative time-discretization for stiff Hamiltonian systems, and molecular chain models

Brenton LeMesurier
Department of Mathematics, College of Charleston, South Carolina
http://blogs.cofc.edu/lemesurierb/

CSCAMM Seminar Wednesday September 26, 2012 (slightly revised the next day)

Contents

1 Introduction

The main objective of this talk is to describe a method for constructing time discretization methods for Hamiltonian systems that

- conserve the Hamiltonian ("energy") and all quadratic and linear conserved quantities, related to affine symmetries,
- respect time reversal symmetry of the equations,
- are unconditionally stable (A-stable for the case of linear systems), and
- are of any chosen (even) order of accuracy.

This is then applied to several "lattice" equations that arise in modeling large molecules and optical waveguide arrays. The method is also applicable to numerous other lattice equations including spatial discretizations of various dispersive nonlinear PDE's, in particular ones of a general nonlinear Schrödinger type.

Some Notation for Derivatives

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

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

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 Some Equations of Interest

2.1 FPU and DNLS

Probably the two most famous example of lattice differential equations are the FPU equation, modeling coupled nonlinear oscillations

$$m_0 \frac{d^2 q_n}{dt^2} = V'(q_{n+1} - q_n) - V'(q_n - q_{n-1}) \tag{1}$$

and the Discrete Nonlinear Schrödinger Equation [DNLS]

$$\mathrm{i} \frac{du_n}{dt} + K(u_{n-1} + u_{n+1}) + 2|u_n|^2 u_n = 0. \tag{2}$$

These are associated respectively with classical mechanical particle systems and with reductions of field equations (including but certainly not limited to the Schrödinger equation) in the presence periodic spatial structures.

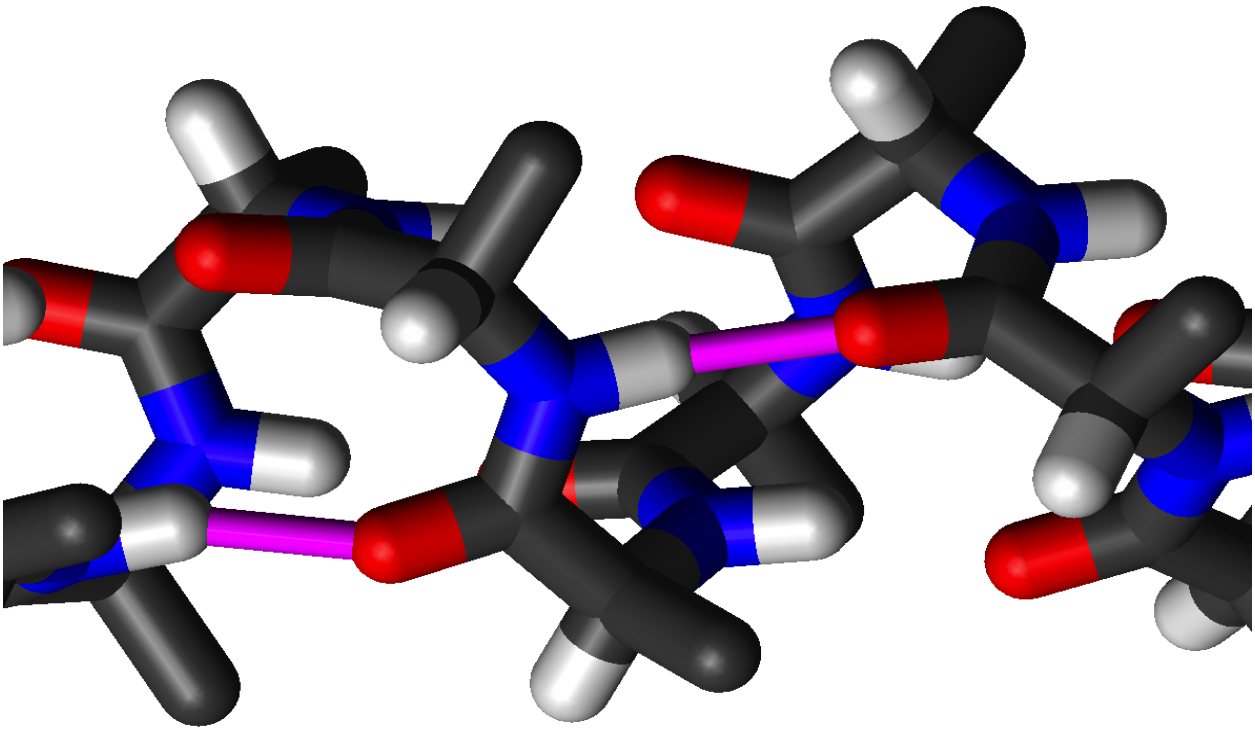

Figure 1: α-helix protein (poly-alanine) stick model. (Source: wikipedia)

2.2 The Davydov System, and Some Reductions

Davydov's Model of Excitation Propagation in Protein

A less familiar equation, combining aspects of the above pair, is Davydov's model of excitation propagation in α-helix protein [Davydov 1971, Davydov&Kislukha 1973]

$$\mathrm{i}\frac{du_n}{dt} + K(u_{n-3} + u_{n+3}) - L(u_{n-1} + u_{n+1}) = (q_{n+3} - q_{n-3})u_n,$$
$$m_0\frac{d^2q_n}{dt^2} - (q_{n-3} - 2q_n + q_{n+3}) = |u_{n+3}|^2 - |u_{n-3}|^2$$

with three physical time scales

$$K \approx 1.4 \text{ THz}, \quad L \approx 2.3 \text{ THz}, \quad \omega_0 = \sqrt{1/m_0} \approx 12 \text{ THz}.$$

This serves as an example of a larger class of so called *exciton-oscillator systems*, and introduces additional challenges to the design of a good numerical methods.

The Variables in Davydov's Model

- The *exciton* variable u_n comes from Schrödinger's equation, and gives the probability that the C=O double bond at the n-th amino acid residue is in a stretching excited state. (This can be reached with the energy quantum of the ATP-ADP interaction, and is 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, in the direction of the axis of the helix.
- The K terms relate to the attractive interaction between excitons in residues that are adjacent along an almost straight *spine* of residues within a helical molecular structure:
there are approximately three residues per twist of the helix.
- The L terms relate to the repulsive interaction between excitons in residues that are adjacent along the molecular backbone of residues within a helical molecular structure.
- The remaining "mechanical" terms relate to the attractive electrostatic dipole interaction between resides adjacent along spines.

Hamiltonian Form for the Davydov System

The Davydov System has a Hamiltonian form, with

$$\begin{aligned}\mathcal{H} = = &-K\sum_n (u_n u_{n+3}^* + u_{n+3}u_n^*) + L\sum_n (u_n u_{n+1}^* + u_{n+1}u_n^*) \\ &+ \sum_n \left[\frac{p_n^2}{2m_0} + \frac{1}{2}(q_{n+3}-q_n)^2\right] \\ &+ \sum_n (q_{n+3}-q_{n-3})u_n u_n^*.\end{aligned} \tag{3}$$

Due to the mix of real and complex variables, the Hamiltonian equations are also most conveniently given as a mix of two types as described next.

Canonical Real Hamiltonian Form

For the real "particle" variables q_n and conjugate momenta $p_n = m_0 q_n$, the equations take the familiar form

$$\frac{dq_n}{dt} = \frac{\partial \mathcal{H}}{\partial p_n}, \quad \frac{dp_n}{dt} = -\frac{\partial \mathcal{H}}{\partial q_n}, \tag{4}$$

or

$$\frac{d\mathbf{q}}{dt} = \nabla_{\mathbf{p}}\mathcal{H}, \quad \frac{d\mathbf{p}}{dt} = -\nabla_{\mathbf{q}}\mathcal{H}. \tag{5}$$

Canonical Complex Hamiltonian Form

For the complex "field" variables u_n, it is instead convenient to use the complex form

$$\frac{du_n}{dt} = \frac{\partial \mathcal{H}}{\partial u_n^*}, \quad \frac{du_n^*}{dt} = -\frac{\partial \mathcal{H}}{\partial u_n}, \tag{6}$$

or

$$\frac{d\mathbf{u}}{dt} = \mathrm{i}\nabla_{\mathbf{u}^*}\mathcal{H}, \quad \frac{d\mathbf{u}^*}{dt} = -\mathrm{i}\nabla_{\mathbf{u}}\mathcal{H}. \tag{7}$$

However the complex form could be eliminated with the change of variables

$$\mathbf{q} = \frac{\mathbf{u}+\mathbf{u}^*}{\sqrt{2}}, \quad \mathbf{p} = \frac{\mathbf{u}-\mathbf{u}^*}{i\sqrt{2}}. \tag{8}$$

Note that $\mathbf{u}$ and $\mathbf{u}^*$ are formally independent variables, though in practice they are complex conjugates, the Hamiltonian is real in that case, and so only the first equation in (7) is needed.

General Hamiltonian Form

It is thus convenient to phrase everything in terms of a slightly more general Hamiltonian form

$$\frac{d\mathbf{y}}{dt} = \mathcal{J}\nabla_{\mathbf{y}}\mathcal{H}(\mathbf{y}) = \mathcal{J}\frac{\partial \mathcal{H}}{\partial \mathbf{y}}(\mathbf{y}) \tag{9}$$

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, $\mathcal{J}(\mathbf{y})$.

Boundary Conditions

Various conditions can be imposed on the Hamiltonian to deal with out-of-bounds indices that arise.

Defining the *bond-stretchings* $\Delta_n := q_{n+3} - q_n$,

- Homogeneous Dirichlet end conditions are most natural physically:

  $$u_n = 0, \qquad \Delta_n := q_{n+3} - q_n = 0$$

  for "out of bounds" values of the index n.

- Periodic end conditions are convenient for testing performance in long-time calculations:

  $$u_{n+N} = u_n, \qquad \Delta_{n+N} = \Delta_n.$$

For constructing PDE approximations via continuum limits, it is also convenient to consider an infinite chain with $n \in \mathbb{N}$ and

$$u_n \to 0, \quad \Delta_n \to 0 \quad \text{as } |n| \to \infty.$$

The Small-and-Fast Oscillation Limit: HDNLS

Many phenomena of interest on the Davydov system are retained in the approximation that the greater stiffness of the mechanical couplings causes the exciton quantities u_n to interact primarily with their time average, given by the singular limit $m_0 \to 0$.

This gives what I will call

The Helical Discrete Schrödinger Equation [HDNLS]

$$\begin{aligned} \mathrm{i}\frac{du_n}{dt} &+ K(u_{n+3} + u_{n-3}) - L(u_{n+1} + u_{n-1}) \\ &+ \left(|u_n|^2 + \frac{1}{2}\left(|u_{n-3}|^2 + |u_{n+3}|^2\right)\right) u_n = 0. \end{aligned} \tag{10}$$

Hamiltonian for HDNLS

The HDNLS has canonical complex Hamiltonian form with

$$\begin{aligned} \mathcal{H} = &- K\sum_n (u_n^* u_{n+3} + u_n u_{n+3}^*) + L\sum_n (u_n u_{n+1}^* + u_{n+1}u_n^*) \\ &+ \frac{1}{2}\sum_n (u_n^* u_n)\left(u_n^* u_n + u_{n+1}^* u_{n+1}\right). \end{aligned} \tag{11}$$

Some Features of HDNLS

Two distinguishing features relative to the standard Discrete Nonlinear Schrödinger [DNLS] equation

$$\mathrm{i}\frac{du_n}{dt} + K(u_{n+1} + u_{n-1}) + 2|u_n|^2 u_n = 0 \tag{12}$$

are the non-local form of the nonlinearity, and the addition of non-nearest neighbor interactions, adding some conformational information about the molecular chain.

A Discrete Nonlinear Schrödinger Equation with a Non-local Nonlinearity

A further simplification for testing purposes is the single spine version of the above,

The Non-Local Discrete Nonlinear Schrödinger Equation [NLDNLS]

$$\begin{aligned} &\mathrm{i}\frac{du_n}{dt} + K(u_{n+1} + u_{n-1}) \\ &\qquad + \left[|u_n|^2 + \frac{1}{2}\left(|u_{n-1}|^2 + |u_{n+1}|^2\right)\right] u_n = 0. \end{aligned} \tag{13}$$

with Hamiltonian

$$\mathcal{H} = -K\sum_n (u_n^* u_{n+1} + u_n u_{n+1}^*) + \frac{1}{2}\sum_n (u_n^* u_n)(u_n^* u_n + u_{n+1}^* u_{n+1}). \tag{14}$$

These exciton-oscillator chain models are interesting challenges for developing and testing numerical methods:

- They cannot be handled by some earlier approaches developed for discretizations of PDEs such as the methods of [LaBudde&Greenspan 1976] or [Strauss&Vázquez 1978], due to the non-local coupling in the nonlinear term, and
- They cannot be handled well by some other popular methods for Lagrangian and mechanical systems, such as the Störmer-Verlet method (which is leap-frog for the case of mechanical systems in standard form).

Non-Mechanical Form

Note that for all the equations with complex field components, the Hamiltonian

1. is not in the standard form for a mechanical system, and indeed
2. does not separate as $\mathcal{H} = T(\mathbf{p}) + U(\mathbf{q})$.

Many popular conservative and symplectic numerical methods that work well for mechanical systems, such as molecular models with no quantum excitations present, rely on the latter splitting.

2.3 Invariants (a.k.a Conserved Quantities, First Integrals)

Conserved "Charge"

The field equations 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 notably, it is quadratic :

$$\mathcal{E} = \sum_n u_n u_n^*. \tag{15}$$

This is associated via Noether's Theorem with a linear symmetry group action, the gauge symmetry

$$\mathbf{u} \to \mathrm{e}^{\mathrm{i}s}\mathbf{u}, \quad \mathbf{u}^* \to \mathrm{e}^{-\mathrm{i}s}\mathbf{u}^*. \tag{16}$$

Conserved Momentum for the Davydov System

The Davydov systems also has a conserved momentum $\mathcal{P}$, again (degenerately) quadratic:

$$\mathcal{P} = \sum_n p_n. \tag{17}$$

This is associated via Noether's Theorem with the symmetry group action $q_n \to q_n + s$.

Conservation of Energy in Hamiltonian Systems

For future reference, here are the verifications of the invariance of the conservation of these invariants.

First, for any system $d\mathbf{y}/dt = \mathcal{J}\nabla_{\mathbf{y}}\mathcal{H}(\mathbf{y})$

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

Phase Shift Invariance and its Invariant Quadratic Forms

Conservation of the charge can be verified directly (rather than invoking Noether's Theorem) by first noting that the state variables u_n and u_n^* appear in the Hamiltonian only through the *symmetry group invariant quadratic combinations*

$$\pi_{nm} = u_n u_m^*$$

which are invariant under the gauge symmetry

$$\mathbf{u} \to \mathrm{e}^{\mathrm{i}s}\mathbf{u}, \ \mathbf{u}^* \to \mathrm{e}^{-\mathrm{i}s}\mathbf{u}^*.$$

Invariance of Charge

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

$$\begin{aligned}\frac{d\mathcal{E}}{dt} &= \sum_n \frac{d(u_n u_n^*)}{dt} \\ &= \sum_n \frac{du_n}{dt}u_n^* + \frac{du_n^*}{dt}u_n \\ &= i\sum_n \frac{\partial \mathcal{H}}{\partial u_n^*}u_n^* - \frac{\partial \mathcal{H}}{\partial u_n}u_n \\ &= i\sum_n \Big[\sum_m \frac{\partial \mathcal{H}}{\partial \pi_{mn}}\frac{\partial \pi_{m,n}}{\partial u_n^*}u_n^* - \sum_m \frac{\partial \mathcal{H}}{\partial \pi_{nm}}\frac{\partial \pi_{nm}}{\partial u_n}u_n\Big] \\ &= i\sum_n \sum_m \Big[\frac{\partial \mathcal{H}}{\partial \pi_{mn}}u_m u_n^* - \frac{\partial \mathcal{H}}{\partial \pi_{nm}}u_n^* u_m\Big].\end{aligned}$$

Then note that for any pair of indices a, b with $a \le b$

- the term $\dfrac{\partial \mathcal{H}}{\partial \pi_{nm}}u_n u_m^*$ for $m = a$, $n = b$ is $\dfrac{\partial \mathcal{H}}{\partial \pi_{ba}}u_b u_a^*$, and
- the term $-\dfrac{\partial \mathcal{H}}{\partial \pi_{mn}}u_n^* u_m$ for $m = b$, $n = a$ is $-\dfrac{\partial \mathcal{H}}{\partial \pi_{ba}}u_a^* u_b$.

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

2.4 More General Lattice Equations

The nonlocal nonlinearity of NLDNLS arises from a "fast mechanical oscillation, slow field excitation" approximation. Similar approximations arise in PDE models, such as with coupling of field propagation to fast acoustic waves in the propagation medium, and lead to the coupling of time evolution equations with time-independent (elliptic) equations. Then solving the time-independent equations to eliminate the fast (mechanical) variables in terms of the slower (field) variables again leads to nonlocal coupling in the nonlinear interaction.

Spatial discretization of such equations, or approximations describing periodic lattice structures, lead to interest in a general class of equations given by Hamiltonians

$$\mathcal{H} = \sum_n \sum_m (l_{nm}u_n u_m^*) + G(\mathbf{u}\mathbf{u}^*)$$

for symmetric matrix $L = \{l_{nm}\}$. Here $\mathbf{u}\mathbf{u}^*$ denotes the matrix of all products $u_n u_m^*$.

The ODEs are then a general family of

Lattice Nonlinear Schrödinger Equations

$$\mathrm{i}\frac{du_n}{dt} + \sum_m l_{nm}u_m + \sum_m g_{nm}(\mathbf{u}, \mathbf{u}^*)u_m = 0. \tag{18}$$

2.5 Features of The Lattice ODE Systems

The systems of ODE's arising from such discretizations of PDEs typically have the following features:

1. A very large number of unknowns, N.
2. A property analogous to quasi-linearity, 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 small in that the Jacobian matrix $D_{\mathbf{y}}\mathbf{g}(\mathbf{y})$ has far smaller norm than L for relevant values of $\mathbf{y}$.
3. The coupling between unknowns is *local*, or *dominantly 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* $w \ll N$.
   (Or cyclically banded in the case of periodic boundary conditions.)
   - *dominantly local* means that the system is quasi-linear as above and the linear part L is banded.

3 Time Discretization Methods That Conserve Energy and "Momenta", and Respect Time Reversal Symmetry

Definition 1 (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.

Methods That Conserve Energy and "Momenta"

An elegant approach to deriving methods that exactly conserve energy (the Hamiltonian) is to approximate the gradients appearing in the Hamiltonian equation

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

This can be done easily, and in many ways, but

- conserving other invariants (often called *momenta*) requires an appropriate choice of the gradient approximation, and
- there is a natural limitation to quadratic (including linear) momenta.

Some Notation for Difference Schemes

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 $\mathbf{u}$:

- $h = \delta t$ 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.
- $t^+ = t + h$ and x^+ denotes the value $x(t^+) = x(t + h)$.
- $\delta x = x^+ - x$.
- $X = \overline{x} = \dfrac{x + x^+}{2}$.
- $\dot{X} = \dfrac{\delta x}{\delta t}$.

Note that capital letters are only used for time averaged or time-step approximated quantities.

3.1 Discrete Gradient Methods

The basic idea, originating in the work of [Gonzales 1996, Gonzales&Simo 1996] is to define a discrete approximation of the gradient

$$\nabla_{\mathbf{y}} f(\mathbf{y}) \approx (\tilde{\nabla}_{\mathbf{y}} f)(\mathbf{y}, \mathbf{y}^+) = \left\langle (\tilde{D}_1 f)(\mathbf{y}, \mathbf{y}^+), (\tilde{D}_2 f)(\mathbf{y}, \mathbf{y}^+), \dots \right\rangle, \tag{19}$$

giving the *discrete Hamilton's equation*

$$\frac{\delta \mathbf{y}}{\delta t} = \mathcal{J}(\tilde{\nabla}_{\mathbf{y}} \mathcal{H})(\mathbf{y}, \mathbf{y}^+). \tag{20}$$

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 [LeM 2012a, LeM 2012b].

From now on we will assume the natural consistency condition

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

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

$$\delta f = (\tilde{\nabla}_{\mathbf{y}} f)(\mathbf{y}, \mathbf{y}^+) \cdot \delta \mathbf{y}, \quad = \sum_n (\tilde{D}_n f)(\mathbf{y}, \mathbf{y}^+) \delta y_n. \tag{21}$$

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 With DIscrete Gradient Methods

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

$$\begin{aligned}\frac{\delta \mathcal{H}}{\delta t} &= (\tilde{\nabla}_{\mathbf{y}} \mathcal{H})(\mathbf{y}, \mathbf{y}^+) \cdot \frac{\delta \mathbf{y}}{\delta t} \\ &= (\tilde{\nabla}_{\mathbf{y}} \mathcal{H})(\mathbf{y}, \mathbf{y}^+) \cdot \mathcal{J} (\tilde{\nabla}_{\mathbf{y}} \mathcal{H})(\mathbf{y}, \mathbf{y}^+) \\ &= 0.\end{aligned}$$

Choosing a Discrete Gradient

Unfortunately, there are infinitely many choices of discrete gradient satisfying this chain rule condition: generically, if you choose all but one component, then the above discrete chain rule determines the value of that last component.

The challenge is to choose a discrete gradient 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,
as was just done for conservation of energy.

Choosing a Discrete Gradient: Respecting Quadratic Invariants

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 factors multiple products in terms of suitable quadratic and linear terms, and this leads again to the result that
quadratic and linear invariants can be handled, but not in general invariants of other forms.

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

Difference Calculus for Functions of One Variable

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

$$\tilde{D}_x f(x, x^+) := \begin{cases} \dfrac{\delta f}{\delta x}, & x^+ \neq x \\ \dfrac{df}{dx}(x), & x^+ = x. \end{cases} \tag{22}$$

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, *symmetric step composition methods* can be used to construct higher order methods, as will be discussed soon.

Avoiding Division By Zero (optional!)

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:

$$\tilde{\nabla}_{\mathbf{y}}(f \circ g)(\mathbf{x},\mathbf{x}^+) = \tilde{D}_g f(g,g^+)\tilde{\nabla}_{\mathbf{x}} g(\mathbf{x},\mathbf{x}^+). \tag{23}$$

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.

Next, we adopt a product rule based on the decomposition $\delta(fg) = \overline{g}\delta f + \overline{f}\delta g$.

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

Time-Symmetric Discrete Product Rule

$$(\tilde{\nabla}_{\mathbf{x}}(fg))(\mathbf{x},\mathbf{x}^+) = \overline{g}(\tilde{\nabla}_{\mathbf{x}} f)(\mathbf{x},\mathbf{x}^+) + \overline{f}(\tilde{\nabla}_{\mathbf{x}} g)(\mathbf{x},\mathbf{x}^+). \tag{24}$$

This is the only choice that has time-reversal symmetry and respects commutativity of products, and it also has the advantage of giving second order accuracy.

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

For linear systems, the Hamiltonian is quadratic and so the above choices lead to the implicit midpoint method. Thus, any choice of discrete gradient method following these rules will have the unconditional A-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 — in particular, 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.

When does this work?

Theorem 2. *If a Hamiltonian system* (9)

(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

$$\mathcal{H}(\mathbf{y}) = \hat{\mathcal{H}}(\hat{\mathbf{y}}(\mathbf{y})) \tag{25}$$

then we get a time discretization that conserves these invariants by using the discrete gradient given by applying the discrete chain rule (23) *to formula* (25)*.*

A Conservative Discrete Gradient Scheme

The resulting time-stepping method is

$$\frac{\mathbf{y}^+ - \mathbf{y}}{h} = \mathcal{J}(\tilde{\nabla}_{\hat{\mathbf{y}}}\hat{\mathcal{H}})(\hat{\mathbf{y}}, \hat{\mathbf{y}}^+) \cdot (\tilde{D}_{\mathbf{y}}\hat{\mathbf{y}})(\mathbf{y}, \mathbf{y}^+). \tag{26}$$

Note:

- The choice of the discrete gradient $(\tilde{\nabla}_{\hat{\mathbf{y}}}\hat{\mathcal{H}})(\hat{\mathbf{y}}, \hat{\mathbf{y}}^+)$ 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 and linearity, through terms like
$$\tilde{D}_{y_a}(y_a y_b) = \bar{y}_b \quad (a \neq b), \quad \tilde{D}_{y_a}((y_a)^2) = 2\bar{y}_a.$$
Thus the formula for $(\tilde{D}_{\mathbf{y}}\hat{\mathbf{y}})(\mathbf{y}, \mathbf{y}^+)$ comes from the formula for the exact Jacobian $\tilde{D}_{\mathbf{y}}\hat{\mathbf{y}}(\mathbf{y})$ 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,

$$\frac{dQ}{dt} = D_{\mathbf{y}}Q(\mathbf{y}) \cdot \mathcal{J} D_{\hat{\mathbf{y}}}\hat{\mathcal{H}} \cdot D_{\mathbf{y}}\hat{\mathbf{y}} = 0 \tag{27}$$

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:

$$\frac{\delta Q}{\delta t} = 0,$$

so $Q(\mathbf{y})$ is conserved by the time-discrete system.

3.2 Practical Implementation: an Iterative Solution Method

A Linearly Implicit Iterative Scheme for Solving Discrete Gradient Systems

The system of equations 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 $\mathbf{y}^{(0)} = \mathbf{y}$ (or some other suitable approximation of $\mathbf{y}^+$).
- Construct successive approximations $\mathbf{y}^{(k)}$ of $\mathbf{y}^+$ by solving

$$\mathbf{y}^{(k+1)} - \mathbf{y} = h\mathcal{J}(\tilde{\nabla}_{\hat{\mathbf{y}}}\hat{\mathcal{H}})(\hat{\mathbf{y}}, \hat{\mathbf{y}}^{(k)}) \cdot (\tilde{D}_{\mathbf{y}}\hat{\mathbf{y}})(\mathbf{y}, \mathbf{y}^{(k+1)}). \tag{28}$$

That is, the nonlinear part $\tilde{\nabla}_{\hat{\mathbf{y}}}\hat{\mathcal{H}}$ is approximated using 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 3. *Each iterate $\mathbf{y}^{(k)}$ given by the above iterative scheme conserves all quadratic first integrals that are conserved by the original discrete gradient scheme* (26).

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 $(\tilde{D}_x)f(x, x^+)$ in (22) only apply to the term $(\tilde{\nabla}_{\hat{\mathbf{y}}}\hat{\mathcal{H}})(\hat{\mathbf{y}}, \hat{\mathbf{y}}^{(k)})$ and instead of depending on the unknown $\mathbf{y}^+$, this depends only on the already-known quantities $\mathbf{y}$ and $\mathbf{y}^{(k)}$.

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.)

3.3 Higher Order Accuracy by Symmetric Step Composition

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 the Hamiltonian and quadratic invariants, and time-reversal symmetry.

The idea is to start with a basic symmetric time-stepping method and construct a composite step from a sequence of s basic steps, of lengths

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

Conditions on the Step Size Factors β_i

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 $p + 2$.

Additional conditions can be found to impose any desired even order of accuracy.

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 $\beta_1 \approx 0.41$, $\beta_2 \approx -0.66$.

3.4 Aside: A Multi-gradient Method for Conserving All Invariants

Multi-gradient System Form

[McLaughlin et al 1999] 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

$$\frac{dy_n}{dt} = \sum_{n_1,\ldots,n_r} A_{n,n_1,n_2,\ldots,n_r} D_{n_1} I_1(\mathbf{y}) \ldots D_{n_r} I_r(\mathbf{y}) \tag{29}$$

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.

4 The Main Alternative: Symplectic Methods

Symplectic Methods: Discretizing Hamilton's Principle

It is natural to compare the methods described above to the more famous *symplectic* methods, which can also give conservation of quadratic invariants, and under suitable conditions are close to conserving energy.

For Hamiltonian $\mathcal{H}(\mathbf{q},\mathbf{p})$ a symplectic method is one with time advance map $(\mathbf{q},\mathbf{p}) \to (\mathbf{q}^+,\mathbf{p}^+)$ given implicitly by the equations

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

where S_h is the generating function ensuring that the map is symplectic.

Symplectic Methods as Discretizations of Hamilton's Principle

This can be interpreted as a discrete approximation of the part of the action interval over time sub-interval t to $t+h$ in the Lagrangian form of the equations: $S = \mathcal{L}_h \approx \int_t^{t+h} \mathcal{L}(\mathbf{q}(s),\dot{\mathbf{q}}(s))\,ds$.

Thus, symplectic methods can be constructed as a discretization of the Lagrangian form of the equations, via a discrete version of the Hamilton's Principle given by the Cauchy-Euler equations for minimization of the discretized action

$$\mathcal{S}_h\left(\left\{\mathbf{q}_\tau\right\}_{\tau=0}^{T}\right) = \sum_{\tau=1}^{T} \mathcal{L}_h(\mathbf{q}_{\tau-1},\mathbf{q}_\tau) \approx \mathcal{S}(\mathbf{q},\dot{\mathbf{q}}) = \int_a^b \mathcal{L}(\mathbf{q}(s),\dot{\mathbf{q}}(s))\,ds.$$

Conserving (Quadratic/Linear) Invariants

If this discretization $\mathcal{L}_h$ is invariant under a continuous symmetry group of the Lagrangian, then the corresponding Noetherian invariants quantities are conserved, by a discrete version of Noether's Theorem [GNI 2006, Section VI.6.4].

(Aside: Many background results are stated here without proof or citation, and in general these will be found in this book Geometric Numerical Integration by E. Hairer, C. Lubich and G. Wanner.)

In practice, this can generally only be true for affine symmetries, and thus for quadratic (including linear) invariants: the same limitation seen for discrete gradient methods.

The most famous and fundamental symplectic method is

The Implicit Midpoint Rule

$$\frac{\mathbf{y}^+ - \mathbf{y}}{h} = f\left(\frac{\mathbf{y} + \mathbf{y}^+}{2}\right) \quad \text{given by } \mathcal{L}_h = \mathcal{L}(\mathbf{Q}, \dot{\mathbf{Q}})\delta t. \tag{30}$$

This is symmetric, unconditionally stable (A-stable), and $\mathcal{L}_h$ respects affine symmetries of $\mathcal{L}$, so that quadratic invariants are conserved.

It also works without reference to the Lagrangian form, convenient for Hamiltonian systems that are not of standard mechanical form $\mathcal{H}(\mathbf{q}, \mathbf{p}) = T(\mathbf{q}) + V(\mathbf{q})$. It has the suggestive form

$$\begin{aligned} \dot{\mathbf{Q}} &= \frac{\partial \mathcal{H}}{\partial \mathbf{p}}(\mathbf{Q}, \mathbf{P}), \\ \dot{\mathbf{P}} &= -\frac{\partial \mathcal{H}}{\partial \mathbf{q}}(\mathbf{Q}, \mathbf{P}). \end{aligned} \tag{31}$$

The Störmer-Verlet Method (not useful here!)

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

This gives

$$\begin{aligned} \mathbf{P} &= \mathbf{p} - \frac{h}{2}\frac{\partial \mathcal{H}}{\partial \mathbf{q}}(\mathbf{P}, \mathbf{q}), \\ \mathbf{q}^+ &= \mathbf{q} + \frac{h}{2}\left(\frac{\partial \mathcal{H}}{\partial \mathbf{p}}(\mathbf{P}, \mathbf{q}) + \frac{\partial \mathcal{H}}{\partial \mathbf{p}}(\mathbf{P}, \mathbf{q}^+)\right), \\ \mathbf{p}^+ &= \mathbf{p} - \frac{h}{2}\frac{\partial \mathcal{H}}{\partial \mathbf{q}}(\mathbf{P}, \mathbf{q}^+), \end{aligned} \tag{32}$$

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

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

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 still explicit:

$$\mathbf{P} = \mathbf{p} - \frac{h}{2}\frac{\partial U}{\partial \mathbf{q}}(\mathbf{q}), \quad \mathbf{q}^+ = \mathbf{q} + h\frac{\partial T}{\partial \mathbf{p}}(\mathbf{P}), \quad \mathbf{p}^+ = \mathbf{p} - \frac{h}{2}\frac{\partial U}{\partial \mathbf{q}}(\mathbf{q}^+).$$

It was for such systems that the method was introduced by C. Störmer in 1903 for astronomical calculations, and independently by L. Verlet in 1967 for molecular dynamics.

However for the more general systems considered here, the Störmer-Verlet Method is less versatile, less stable, and less accurate than the implicit midpoint method, so will not be considered further here.

"Symplectic Methods Cannot Conserve Energy"

Symplectic methods such as the implicit midpoint method cannot conserve the Hamiltonian ("energy"), except in degenerate cases, due to a theorem of [Ge & Marsden 1988].

More precisely, if a time discretization with uniform step size preserves both the symplectic form and the Hamiltonian, then either

- the scheme is the exact time stepping map except for shifts in time, or
- the system decomposes into several parts for which the above is true.

Linear systems are the exception that proves the rule: The energy is quadratic and so is conserved by symplectic methods such as the implicit midpoint rule. The system decomposes through action-angle variables into components with one dimensional orbits that are level curves of quadratic invariants for each component. Note that the implicit midpoint rule and discrete gradient method are identical in this case.

Exponentially Small Errors In Energy

The non-conservation of energy by symplectic methods is mitigated in many cases by theorems which, loosely speaking, guarantee that the error in energy is exponentially small in the time step size h for an exponentially long time.

The exponential term is of the form $\mathrm{e}^{-h_0/h}$ where the time scale h_0 is the inverse of the norm of the Jacobian of the right-hand side of the equation, so the extraordinary accuracy of energy is only realized with step size h significantly smaller than h_0.

Note that h_0 is roughly the time step size limit for

- stability of explicit schemes, and
- convergence of a simple fixed point iterative method for an implicit scheme.

Exponentially Small Errors In Energy: Not Useful For Stiff Systems

However, for stiff systems one typically wants to use step size $h > h_0$, and even $h \gg h_0$.

In this case, current best practice for Hamiltonian systems seems to be *symmetric split step methods,* in which one sub-step is for the stiff, linear modes only, in turn done with many sub-steps or with an A-stable method.

Higher Order Symplectic Runge-Kutta Methods, by Step Composition

One way to produce symplectic methods with higher order of accuracy is to again use step composition. Indeed this is what those step composition methods were first developed for.There is a sense is which this is almost the best possible approach:

Theorem 4 ([GNI 2006], Theorem VI.4.4)**.** *Any [symmetric], symplectic, irreducible, diagonally implicit Runge-Kutta [DIRK] method is given by [symmetric] step composition with the implicit midpoint rule as the basic step.*

Recall that s-stage DIRK methods are of the form

$$\begin{aligned}\frac{\delta y^{(k)}}{\delta t} &= f\left(y + \sum\nolimits_{j=1}^{k} \alpha_{jk}\delta y^{(k)}\right), \quad 1 \le k \le s,\\ \delta y &= \sum\nolimits_{k=1}^{s} \beta_k \delta y^{(k)}.\end{aligned}$$

(The form is given only for autonomous systems, WSLOG.)

DIRK methods require implicit solution, but only for one approximate increment $\delta y^{(k)}$ at a time, in order. Thus there is a correspondence between any such method a a composition of DG steps with the methods identical for the linear parts of the equation.

Comparisons between the basic (second order) methods for equations like DNLS favor the DG method; it remains to be seen how the corresponding higher order methods compare.

Fully Implicit Higher Order Symplectic Runge-Kutta Methods

One other category of symplectic methods is of interest: fully implicit s- stage Runge-Kutta methods

$$\begin{aligned}\frac{\delta y^{(k)}}{\delta t} &= f\left(y + \sum\nolimits_{j=1}^{s} \alpha_{jk}\delta y^{(k)}\right), \quad 1 \le k \le s,\\ \delta y &= \sum\nolimits_{k=1}^{s} \beta_k \delta y^{(k)}\end{aligned}$$

of Gaussian form, related to Gaussian quadrature formulas and, like them, of order $2s$.

The case $s = 1$ gives the Implicit Midpoint Method again, while $s = 2$ give the fourth order Gaussian method

$$\begin{aligned}\frac{\delta y^{(1)}}{\delta t} &= f\left(y + \frac{1}{4}\delta y^{(1)} + \left(\frac{1}{4} - \frac{1}{2\sqrt{3}}\right)\delta y^{(2)}\right)\\ \frac{\delta y^{(2)}}{\delta t} &= f\left(y + \left(\frac{1}{4} + \frac{1}{2\sqrt{3}}\right)\delta y^{(1)} + \frac{1}{4}\delta y^{(2)}\right)\\ \delta y &= \frac{\delta y^{(1)} + \delta y^{(2)}}{2}.\end{aligned}$$

5 Discrete Gradient vs Symplectic Methods, for Stiff Systems

Discrete Gradient vs Symplectic Methods

For general (non-stiff) systems where one wishes to resolve all time scales in the solution accurately, evidence suggests that a Gaussian method solved by fixed point iteration is usually the best approach.

However, this assumes that one is wiling to use time steps $h < h_0$, and further that either one is willing to use step sizes small enough to get extraordinary (exponentially small) errors in energy, or that such extraordinary accuracy is unimportant.

If instead one wishes to solve a stiff system with step size $h > h_0$, so that simple fixed point iteration cannot be used:

1. the cost-accuracy balance move in favor of the DIRK symplectic methods given step composition of the midpoint rule over Gaussian methods, and
2. energy accuracy and empirical observations favor the corresponding step compositions of the discrete gradient method over symplectic DIRK methods.

6 Numerical Results and Partial Explanation

We conclude with a few numerical results, first to test the method, and then make a few observations about continuum limits of DNLS equations.

6.1 Testing With The Non-Locally Nonlinear DNLS

For a comparison of the discrete gradient method to the popular symplectic 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 (13)

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

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

Thus 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.

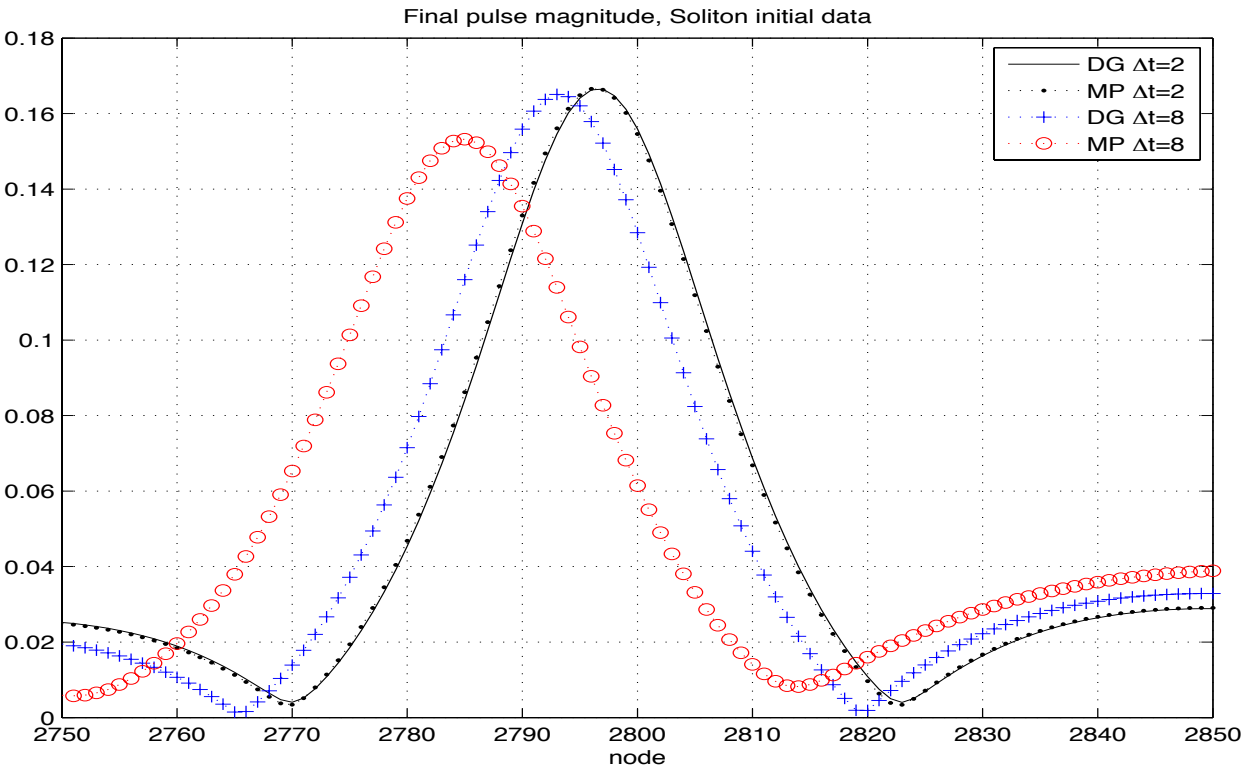

Figure 2: NLDNLS, $A = 2$: $|u_n|$ at time $t = 40000$.

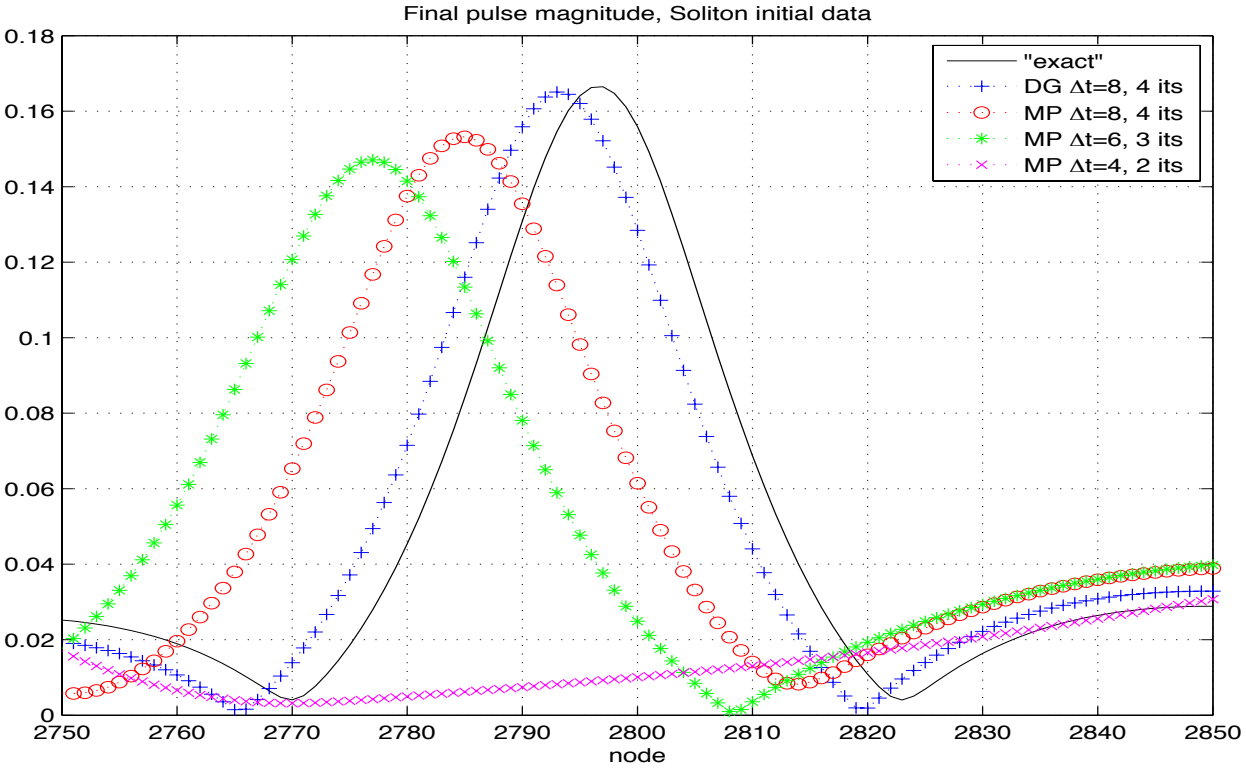

Figure 3: NLDNLS, $A = 2$: $|u_n|$ at time $t = 40000$, methods of roughly equal time cost.

6.2 Numerical Observations on The Helical Discrete Nonlinear Schrödinger Equation

Results for the Helical DNLS System

The final numerical results are for Helical DNLS system with impulsive initial data at one end of the molecular chain,

$$u_1(0) = 1 \text{ and } \quad u_n(0) = 0 \text{ for } n > 1,$$

and more generally with $u_n(0) = 0$ for $n > 2$.

Numerical Observations

Several features are seen in the graphs here and confirmed for a wide range of impulse initial data, initial nonzero on only the first one or two nodes:

1. A leading pulse with slowly varying amplitude, and **speed of approximately 12.8** (nodes per unit time), followed by a decreasing oscillatory tail, **reminiscent of the Airy function** Ai.
2. The tail with slow amplitude variation extends over the front half of the nodes behind which it is interrupted by a second pulse and the end of slow amplitude variation: the the pulse width is growing, but the spatial period of its oscillations varies little, so the number of oscillations increases.
3. However, **the signal is not slowly varying**, because the phase shifts by roughly a factor i between consecutive nodes.
4. This is seen in the near-linear case of very small amplitude, but the proportion of the charge in the pulse increases as total charge (degree of nonlinearity)
5. At even larger amplitudes, the lead hump becomes more dominant and the form moves towards a somewhat sech-like pulse surrounded by far lower amplitude oscillations. However the i^n phase pattern persists, so there is still no NLS continuum limit.
6. Similar behavior is seen in the basic DNLS equation, but with differences such as reversal of the phase factor to $(-\mathrm{i})^n$.

6.3 A Different Continuum Limit for Various DNLS Models

Linear Dispersion Relation and Signal Speed

The explanation of these phenomena seems to come from the linearization and bifurcation there-from, related to work of [Pelinovsky&Rothos 2005] and [Kevrekidis et al 2007].

First, consider the linear part of the HDNLS equation

$$\mathrm{i}\frac{du_n}{dt} + K(u_{n+3} + u_{n-3}) - L(u_{n+1} + u_{n-1}) = 0$$

and seek travelling waves of the form

$$u_n = \phi(z)\mathrm{e}^{\mathrm{i}\beta n + \omega t}, \quad z = n - vt, \quad \phi(z) = \mathrm{e}^{\mathrm{i}kt}$$

with an emphasis on slow variation in the sense of small values of ω and k. This gives the dispersion relation

$$\omega(k) = kv + 2L\cos(3(k+\beta)) - 2K\cos(k+\beta).$$

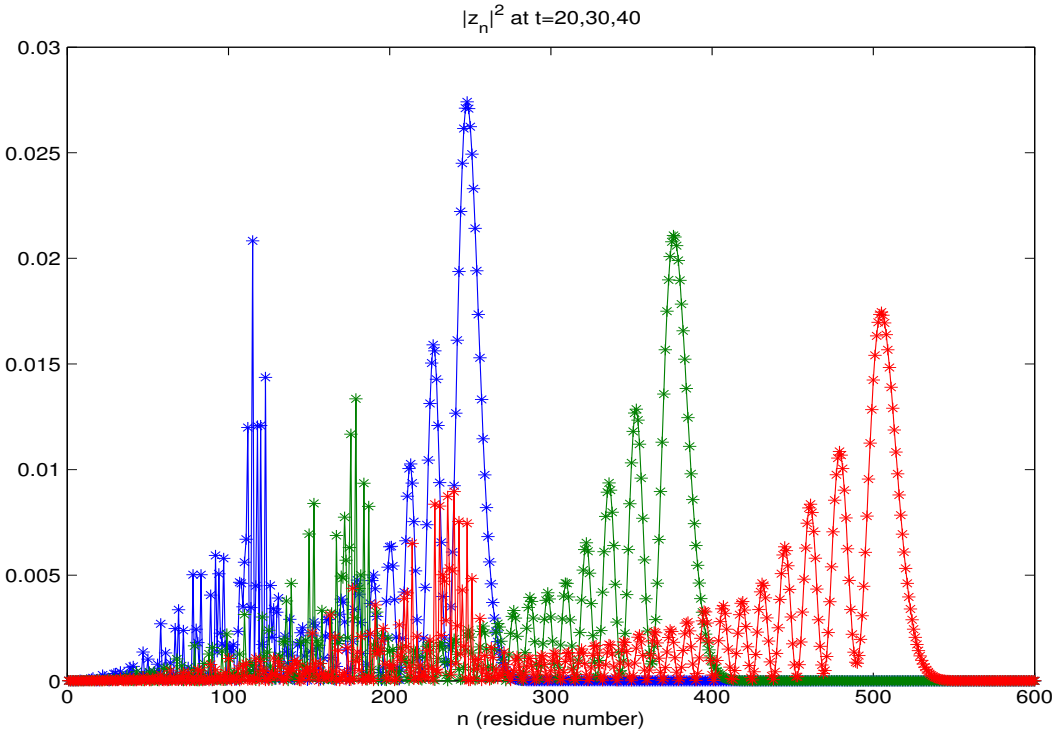

Figure 4: HDNLS: $|u_n|^2$ at time $t = 20, 30, 40$.

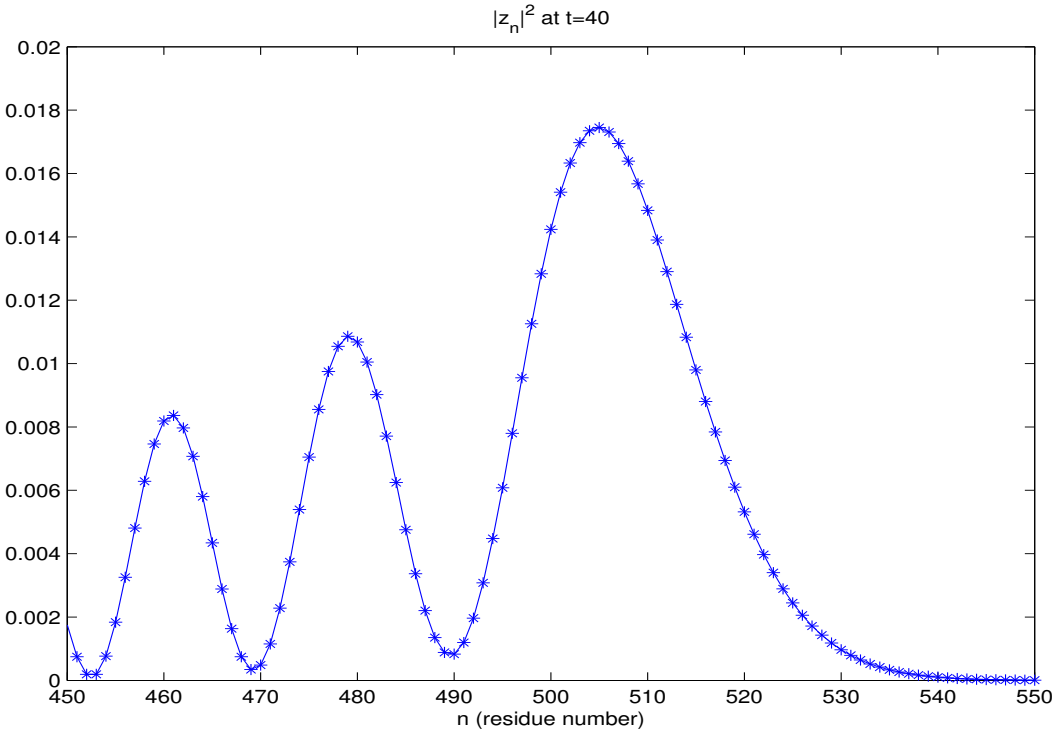

Figure 5: HDNLS: $|u_n|^2$ at time $t = 20, 30, 40$, near main pulse — a continuum limit?

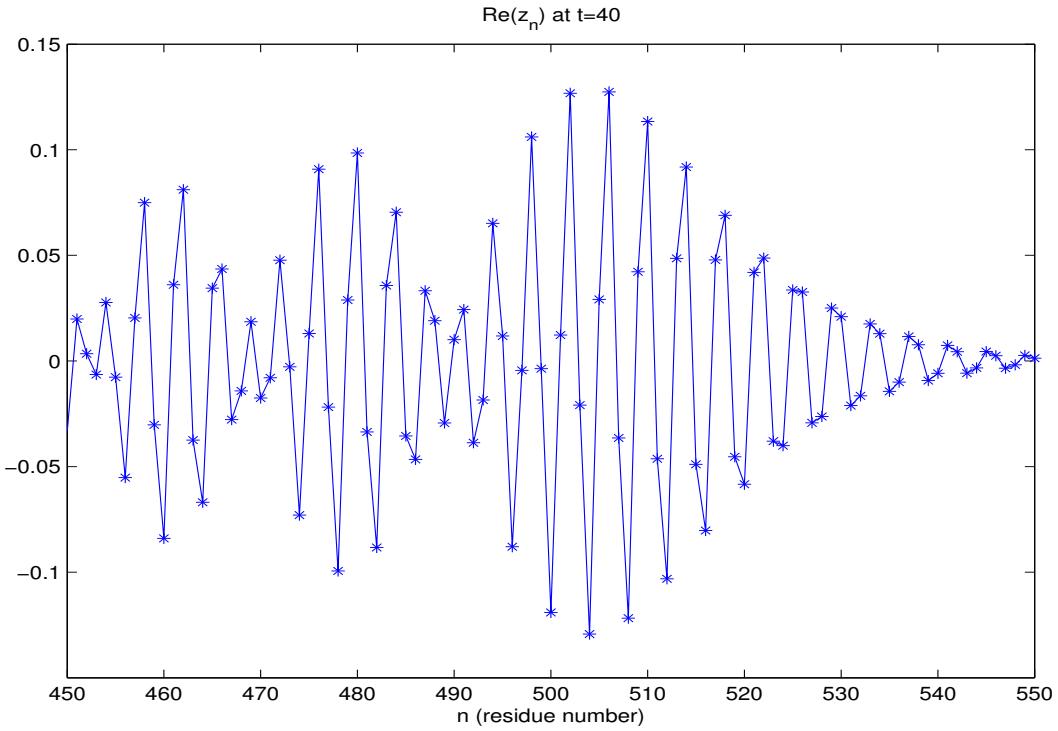

Figure 6: HDNLS: Re(u_n) at time $t = 40$ — no continuum limit!

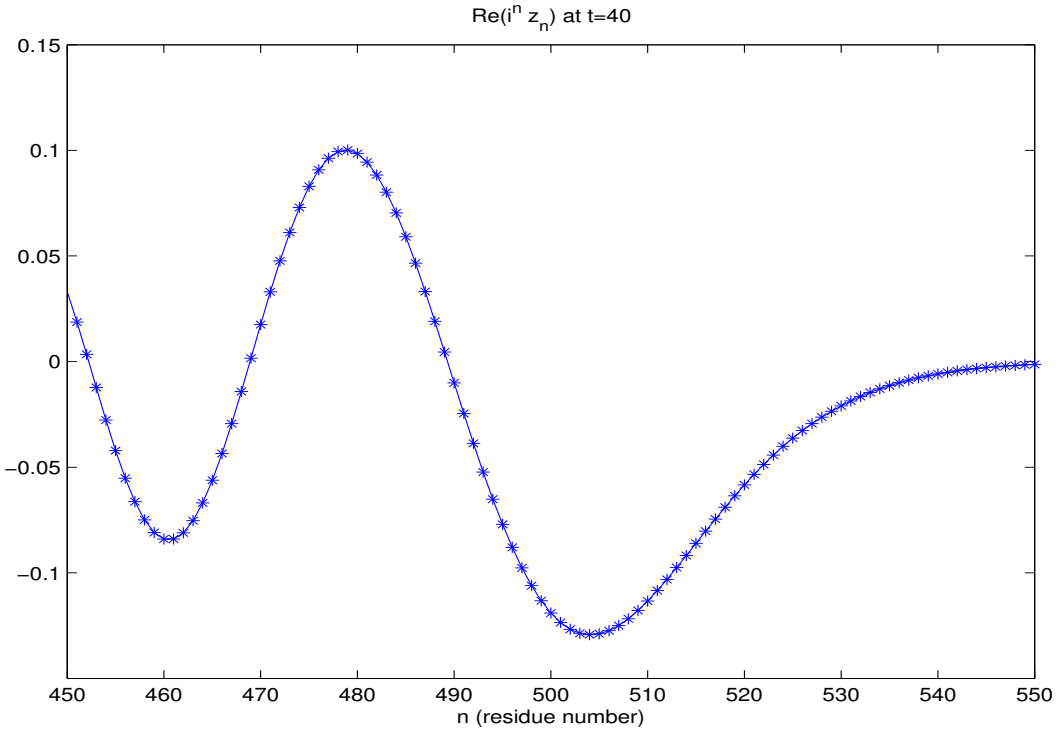

Figure 7: HDNLS: Re$(\mathrm{i}^n u_n)$ at time $t = 40$: — maybe a continuum limit now.

The inflection points at $k = 0, \omega'(0) = 0$ gives group velocity

$$v = 6L\sin(3\beta) - 2K\sin(\beta)$$

with maximum signal speed

$$v_{max} = 6L - 2K \text{ (= 13 for the physical parameter values used here)}$$

occurring for $\beta = -\pi/2$, and thus $\omega = 0$.
This corresponds to small amplitude solutions of the form

$$z_n \approx (-\mathrm{i})^n \phi(z) \mathrm{e}^{\mathrm{i}\omega t},$$

with k and ω small, so that the quantity that varies slowly along the lattice is not z_n but

$$w_n := \mathrm{i}^n z_n.$$

A Different Continuum Limit

A continuum limit expansion in the limit of a small, slowly varying solution $w_n(t) \approx \epsilon w(\epsilon^3 t, z)$ and a few others rescalings to normalize constants gives what is sometimes called *the third derivative nonlinear Schrödinger equation*

$$\frac{\partial w}{\partial t} = \frac{\partial^3 w}{\partial z^3} + \mathrm{i}2|w|^2 w.$$

Linear Approximation: The Airy Diffusion Equation

The linear part is the strangely named "Airy Diffusion Equation", with solution for $u(0, z) = f(z)$ given in terms of the Airy function Ai as

$$w(t, z) = \int_{-\infty}^{\infty} \frac{f(y)}{(3t)^{(1/3)}} \mathrm{Ai}\left[(z - y)(3t)^{(1/3)}\right] dy$$

Thus with suitable impulsive initial data and boundary conditions there is a slowly decaying and spreading solution

$$w(t, z) = \frac{1}{(3t)^{(1/3)}} \mathrm{Ai}\left[(z - y)(3t)^{(1/3)}\right].$$

Approximate Solutions of Discrete System: Airy Function Pulses

Put back in terms of $x = z + vt$ and $w_n(t)$, this corresponds to an approximate solution of the discrete system with

- Airy function profile.
- Approximately traveling wave form, of speed $2L + 6K = 13$.
- Slow decrease in amplitude and broadening, with time scale $t^{1/3}$.

Nonlinear Self-Focusing Effects? Deviation from Airy Function Pulse Form

This does not fit exactly what is seen due to its doubly infinite domain and infinite charge, but it seems to contains some essential elements of the actual solutions seen.

One difference is that the solutions of HDLNS do not appear to spread at rate $t^{1/3}$.

Perhaps this is the familiar control of dispersion by nonlinear focusing. This idea is supported by the numerical observation that for stronger nonlinearity, the lead pulse is higher and narrower, and shows little or no time-decay in amplitude.

7 Some Plans

1. Implement and assess performance of the fourth order discrete gradient methods given by the Suzuki fractal method, and compare to currently preferred methods such as the fourth-order Gauss method the symplectic diagonally implicit Runge-Kutta method given by Suzuki composition of midpoint steps, and split-step methods for stiff Hamiltonian systems.

2. Study other systems, such as 2D lattice equations from models of thin bio-molecular films, and models of multiple-core fiber optics.

3. Add small dispersion while retaining correct conservation of charge and accurate evolution of energy.

4. Add fixed pattern noise and stochastic terms.

5. Find out more about the new continuum limit equation

$$\frac{\partial z}{\partial t} = \frac{\partial^3 z}{\partial x^3} + 2\mathrm{i}|z|^2 z.$$

6. Analyze bifurcation of solutions of HDNLS from the linear plane wave solutions at $k = \omega = 0$.

Bibliography

[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. *Energy and momentum conserving methods of arbitrary order for the numerical integration of the equations of motion. Parts I and II*, 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.

[Scott 1982] Alwyn C. Scott. The vibrational structure of Davydov solitons. *Physica Scripta*, 25:651–658, 1982.

[Ge & Marsden 1988] Z. Ge and J. E. Marsden. *Lie-Poisson Hamilton-Jacobi theory and Lie-Poisson integrators*, 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. *Fractal decomposition of exponential operators with applications to many-body theories and Monte-Carlo simulations*, Phys. Lett. A 146 (1990) 319–323.

[Yoshida 1990] H. Yoshida. *Construction of higher order symplectic integrators*, Phys. Lett. A 150 (1990) 262–268.

[Simo et al 1992] Juan C. Simo, N. Tarnow and K. K. Wong. *Exact energy-momentum conserving algorithms and symplectic schemes for nonlinear dynamics*, Comput. Methods Appl. Mech. Eng. 100 (1992) 63–116.

[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.

[Pelinovsky&Rothos 2005] D. Pelinovsky and V. Rothos. *Bifurcations of travelling wave solutions in the discrete NLS equation*, Physica D 202 (2005) 16—36.

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

[LeM 2012a] Brenton LeMesurier. *Conservative unconditionally stable discretization methods for Hamiltonian equations, applied to wave motion in lattice equations modeling protein molecules. Physica D*, 241(1):1–10, January 2012. (Published online 1 Oct 2011. DOI:10.1016/j.physd.2011.09.012).

[LeM 2012b] Brenton LeMesurier. *Studying Davydov's ODE model of wave motion in alpha-helix protein using exactly energy-momentum conserving discretizations for Hamiltonian systems. Mathematics and Computers in Simulation*, 82(7):1239—1248, March 2012. (Published online 30 December 2010 DOI:10.1016/j.matcom.2010.11.017).