Summerschool in Aveiro (Sept. 2018), Ernst Hairer

- **Part I. Geometric numerical integration**
  - Hamiltonian systems, symplectic mappings, geometric integrators, Störmer–Verlet, composition and splitting, variational integrator
  - Backward error analysis, modified Hamiltonian, long-time energy conservation, application to charged particle dynamics

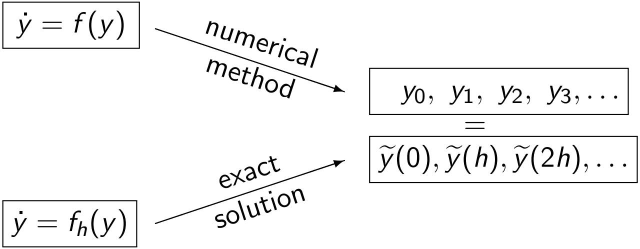
- **Part II. Differential equations with multiple time-scales**
  - Highly oscillatory problems, Fermi–Pasta–Ulam-type problems, trigonometric integrators, adiabatic invariants
  - Modulated Fourier expansion, near-preservation of energy and of adiabatic invariants, application to wave equations

Lecture 2. Backward error analysis

1. **Modified differential equation**
   - Construction of the modified differential equation
   - Hamiltonian systems – symplectic methods
2. **Long-time error analysis**
   - Near-energy preservation
   - Linear error growth for integrable systems
3. **Application to charged particle dynamics**
   - Basic properties
   - Main result - energy preservation
   - Numerical experiments
   - Proof – backward error analysis

Modified differential equation

Given a differential equation $\dot{y} = f(y)$ and a method $y_{n+1} = \Phi_h(y_n)$



Find a *modified differential equation* $\dot{y} = f_h(y)$ of the form

$$\dot{y} = f(y) + h\, f_2(y) + h^2 f_3(y) + h^3 f_4(y) + \ldots$$

such that its solution $\widetilde{y}(t)$ satisfies formally $\quad y_n = \widetilde{y}(nh)$.

Construction of the modified differential equation

Numerical method for $\dot y = f(y)$

$$y_1 = \Phi_h(y_0) = y_0 + hf(y_0) + h^2 d_2(y_0) + h^3 d_3(y_0) + \dots$$

Ansatz for the **modified equation**

$$\dot y = f(y) + hf_2(y) + h^2 f_3(y) + \dots, \qquad y(0) = y_0$$

Taylor series expansion of its solution $\widetilde y(t)$ at $t = h$

$$\begin{aligned}\widetilde y(h) &= y_0 + h\widetilde y\,'(0) + \tfrac{h^2}{2!}\,\widetilde y\,''(0) + \tfrac{h^3}{3!}\,\widetilde y\,'''(0) + \dots \\ &= y_0 + h\big(f + hf_2 + h^2 f_3 + \dots\big)_0 \\ &+ \tfrac{h^2}{2!}\big(f' + hf_2' + \dots\big)\big(f + hf_2 + \dots\big)_0 + \dots \ .\end{aligned}$$

Comparison of like powers of h yields

$$\begin{aligned}d_2(y) &= f_2(y) + \tfrac{1}{2!} f'f(y) \\ d_3(y) &= f_3(y) + \tfrac{1}{2!}\big(f'f_2 + f_2'f\big)(y) + \tfrac{1}{3!}\big(f''(f,f) + f'f'f\big)(y) \\ &\dots\end{aligned}$$

Modified equations for the pendulum equation

explicit Euler:

$$\begin{pmatrix}\dot q\\ \dot p\end{pmatrix} = \begin{pmatrix}p\\ -\sin q\end{pmatrix} + \frac{h}{2}\begin{pmatrix}\sin q\\ p\cos q\end{pmatrix} + \frac{h^2}{12}\begin{pmatrix}-4\,p\cos q\\ (p^2+4\cos q)\sin q\end{pmatrix} + \ldots$$

implicit Euler: same equation with h replaced by $-h$.

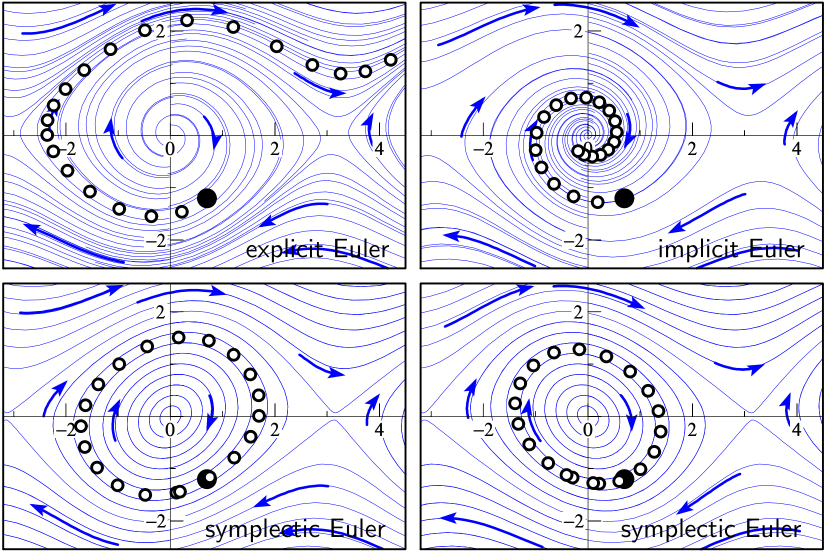
symplectic Euler (explicit in q, implicit in p):

$$\begin{pmatrix}\dot q\\ \dot p\end{pmatrix} = \begin{pmatrix}p\\ -\sin q\end{pmatrix} + \frac{h}{2}\begin{pmatrix}-\sin q\\ p\cos q\end{pmatrix} + \frac{h^2}{12}\begin{pmatrix}2\,p\cos q\\ (p^2-2\cos q)\sin q\end{pmatrix} + \ldots$$

symplectic Euler (explicit in p, implicit in q):

same equation with h replaced by $-h$

Numerical illustration: pendulum $\dot{q} = p, \ \dot{p} = -\sin q$



expl.in p, impl. in q | expl.in q, impl. in p

Quadrature

Problem: $\dot{y} = f(t), \quad y(0) = 0$

Method: $y_{n+1} = y_n + \frac{h}{2}\big(f(t_{n+1}) + f(t_n)\big)$

The modified differential equation is of the form

$$\dot{y} = f(t) + hb_1 f'(t) + h^2 b_2 f''(t) + h^3 b_3 f'''(t) + \ldots$$

The coefficients can be obtained by putting $f(t) = e^t$:

$$(e^h - 1)(1 + hb_1 + h^2 b_2 + \ldots) \;=\; \frac{h}{2}(e^h + 1)$$

This is the generating function for the Bernoulli numbers; we have

$$b_k \approx c\,(2\pi)^{-k}$$

Conclusion. As soon as $f^{(k)}(t) \approx k!\,M\,R^{-k}$, the series of the modified equation diverges for all $h \neq 0$.

Hamiltonian systems

Consider a Hamiltonian system

$$\dot{y} = J^{-1} \nabla H(y)$$

and a one-step method

$$y_{n+1} = \Phi_h(y_n)$$

What can be said about its modified differential equation?

Example (pendulum)
explicit and implicit Euler: modified equation is **not** Hamiltonian
symplectic Euler: modified equation is Hamiltonian with

$$H_h(p,q) = \frac{1}{2}\,p^2 - \cos q - \frac{h}{2}\,p\,\sin q + \frac{h^2}{12}(p^2 - \cos q)\cos q + \ldots$$

Is this true in general?

Modified equation for symplectic methods

Theorem

Consider

- *a Hamilton system with smooth* $H : U \to \mathbb{R}$
- *a symplectic integrator* $\Phi_h(y)$

Then, the vector fields $f_k(y)$ *of the modified differential equation are Hamiltonian, i.e., we have* $f_k(y) = J^{-1}\nabla H_k(y)$.

There are several proofs for this result.

Local existence of the Hamiltonian $H_k(y)$

simple proof by induction without additional assumption
idea goes back to Moser (1968),
Benettin & Giorgilli (1994), Tang (1994)

Global existence of the Hamiltonian $H_k(y)$

needs additional assumptions (satisfied by RK methods)
proof via generating functions: Murua (1994)
algebraic proof for B-series integrators: Hairer (1994)

Proof by induction (local existence)

We prove by induction on N that

$$\dot{y} = f(y) + hf_2(y) + \ldots + h^{N-1} f_N(y) \qquad (1)$$

is Hamiltonian. This is obviously true for $N = 1$.
Assume that (1) is Hamiltonian for N. Its flow $\varphi_{N,t}(y)$ satisfies

$$\Phi_h(y) = \varphi_{N,h}(y) + h^{N+1} f_{N+1}(y) + \mathcal{O}(h^{N+2}).$$

Since $\Phi_h(y)$ and $\varphi_{N,t}(y)$ are symplectic, it holds

$$\begin{aligned} J &= \Phi_h'(y)^{\mathsf{T}} J \, \Phi_h'(y) = \ldots \\ &= J + h^{N+1}\big(f_{N+1}'(y)^{\mathsf{T}} J + J \, f_{N+1}'(y)\big) + \mathcal{O}(h^{N+2}), \end{aligned}$$

so that $J \, f_{N+1}'(y)$ is symmetric. The integrability lemma implies the local existence of $H_{N+1}(y)$ such that $J \, f_{N+1}(y) = \nabla H_{N+1}(y)$. □

Modified equation for symmetric methods

Theorem (adjoint method)

a) Let $f_j(y)$ *be the coefficient functions of the modified differential equation for a method* $\Phi_h(y)$*. The coefficient functions of the modified equation for the adjoint method* $\Phi_h^*(y) = \Phi_{-h}^{-1}(y)$ *are then given by*

$$f_j^*(y) = (-1)^{j+1} f_j(y).$$

b) The modified equation of a symmetric method has an expansion in even powers of h.

Proof.

The solution $\widetilde{y}(t)$ of the modified equation for Φ_h^* has to satisfy

$$\widetilde{y}(t) = \Phi_{-h}\big(\widetilde{y}(t+h)\big) \qquad \text{and hence} \qquad \widetilde{y}(t-h) = \Phi_{-h}\big(\widetilde{y}(t)\big)$$

Replacing h by $-h$, we get the solution of the modified equation for the method Φ_h. □

Structure preservation

There are many similar results that can all be proved by the same induction argument.

- divergence-free vector fields $\dot{y} = f(y)$, i.e., $\mathrm{div} f(y) = 0$, and volume-preserving flows,
- Poisson systems $\dot{y} = B(y)\nabla H(y)$ and Poisson mappings,
- vector fields on a manifold and flows on the manifold,
- special case, where the manifold is a Lie group,
- differential equations with first integrals.

Always when a numerical integrator shares a characteristic property of the exact flow, the modified differential equation retains the structure of the problem.

Lecture 3. Backward error analysis

1. **Modified differential equation**
   - Construction of the modified differential equation
   - Hamiltonian systems – symplectic methods

2. **Long-time error analysis**
   - Near-energy preservation
   - Linear error growth for integrable systems

3. **Application to charged particle dynamics**
   - Basic properties
   - Main result - energy preservation
   - Numerical experiments
   - Proof – backward error analysis

Estimates of the local error

Since the modified equation is in general divergent, we have to truncate it. What is the induced error?

Theorem (local error estimation)

Denote by $\varphi_{N,t}(y)$ *the flow of the truncated modified differential equation*

$$\dot{y} = f(y) + hf_2(y) + h^2 f_3(y) + \ldots + h^{N-1} f_N(y),$$

then there exists a constant $C_N(y_0)$ *such that for* $h \leq h_0$

$$\|\Phi_h(y_0) - \varphi_{N,h}(y_0)\| \leq C_N(y_0)\, h^{N+1}.$$

The proof is trivial. One even knows that

$$\Phi_h(y_0) - \varphi_{N,h}(y_0) = h^{N+1} f_{N+1}(y_0) + \mathcal{O}(h^{N+2}).$$

It is less trivial to study the dependence of $C_N(y_0)$ on N.

Exponentially small error estimates

Typically (e.g., when $f(y)$ is real-analytic) one cannot expect a better estimate than

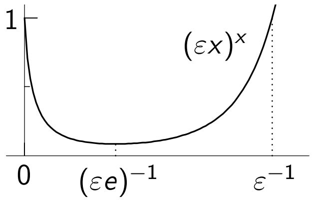
$$C_N(y_0) \leq \alpha(\omega N)^N$$

Optimal choice of N:
the estimate

$$C_N(y_0)h^{N+1} \leq h\alpha(\omega h N)^N$$

is minimal, when

$$N = (\omega h e)^{-1}.$$



This choice of N yields

$$\|\Phi_h(y_0) - \varphi_{N,h}(y_0)\| \leq h\alpha\, e^{-\gamma/h} \quad \text{with} \quad \gamma = (\omega e)^{-1} > 0$$

Estimation of the global error

To get estimates of the global error, one has to know something about the error propagation of the modified differential equation.

- **Typical situation:** if we know that

  $$\|\widetilde{\varphi}_t(y_0) - \widetilde{\varphi}_t(z_0)\| \le c\, e^{\omega t}\, \|y_0 - z_0\|,$$

  then we have for $t = nh$

  $$\|y_n - \widetilde{\varphi}_t(y_0)\| \le \alpha\, e^{-\gamma/h}\, t\, e^{\omega t}$$

  exponentially close on intervals of length $\mathcal{O}(1)$.

- **Integrable systems:** if we know that

  $$\|\widetilde{\varphi}_t(y_0) - \widetilde{\varphi}_t(z_0)\| \le (a + b\, t)\, \|y_0 - z_0\|,$$

  then we have for $t = nh$

  $$\|y_n - \widetilde{\varphi}_t(y_0)\| \le \alpha\, e^{-\gamma/h}\, (at + bt^2/2)$$

  exponentially close on exponentially long time intervals.

Near-energy preservation

Theorem

Consider a symplectic method of order r with global modified Hamiltonian

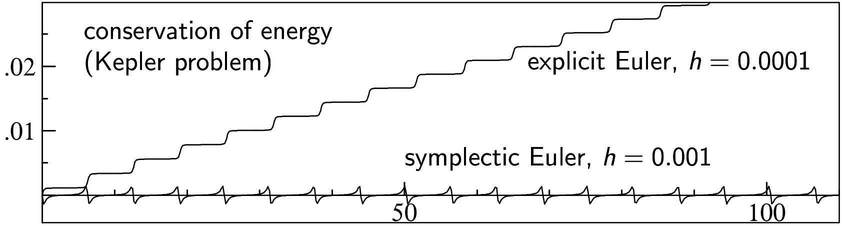
$$\widetilde{H}(p,q) = H(p,q) + h^r H_{r+1}(p,q) + \ldots + h^{N-1} H_N(p,q).$$

Then, the numerical solution satisfies

$$H(p_n,q_n) \;=\; H(p_0,q_0) + \mathcal{O}(h^r) \qquad \text{for} \quad nh \le e^{\gamma/2h}$$

as long as the numerical solution stays in a compact set.

Proof. We have $|\widetilde{H}(p_n,q_n) - \widetilde{H}(p_0,q_0)| \le C\, nh\, e^{-\gamma/h}$. □



Completely integrable systems

Consider a Hamiltonian system

$$\dot{y} = J^{-1} \nabla H(y)$$

Definition

A Hamiltonian system with d degrees of freedom ($H : M \to \mathbb{R}$ with an open set $M \subset \mathbb{R}^d \times \mathbb{R}^d$) is called *completely integrable* if there exist smooth functions $F_1 = H, F_2, \ldots, F_d$ such that

- $F_1, \ldots, F_d$ are in convolution, i.e., $\{F_i, F_j\} = \nabla F_i^{\mathsf{T}} J^{-1} \nabla F_j = 0$,
- The gradients of $F_1, \ldots, F_d$ are everywhere linearly independent,
- The solution trajectories of the Hamiltonian systems with F_i ($i = 1, \ldots, d$) exist for all times and remain in M.

Note that the first condition implies that all function F_j are first integrals (conserve quantities) of the Hamiltonian system.

Completely integrable systems – examples

- Hamiltonian systems with **one** degree of freedom, e.g., harmonic oscillator, mathematical pendulum
- Kepler problem
  first integrals are: energy H and angular momentum $L = q_1 p_2 - q_2 p_1$
- Toda lattice with $H(p,q) = \sum_{k=1}^{n} \Big(\frac{1}{2}\, p_k^2 + \exp(q_k - q_{k+1}) \Big)$
  related to a Lax pair $\dot{L} = [B(L), L]$ with skew-symmetric $B(L)$
- Ablowitz–Ladik discrete nonlinear Schrödinger equation
- Volterra lattices

Arnold–Liouville Theorem

For a *completely integrable* Hamiltonian system there exists a symplectic transformation

$$(p,q) = \psi(a,\theta) \quad (2\pi\text{-periodic in } \theta)$$

to *action-angle variables* such that the Hamiltonian becomes

$$H(p,q) = H(\psi(a,\theta)) = K(a).$$

In the action-angle variables, the system becomes

$$\dot a_i = 0, \quad \dot\theta_i = \omega_i(a), \qquad i = 1,\dots,d$$

with $\omega_i(a) = \partial K/\partial a_i(a)$, and can be solved directly

$$a_i(t) = a_{i0}, \quad \theta_i(t) = \theta_{i0} + \omega_i(a_0)t$$

so that

$$(p(t),q(t)) = \psi(a_0, \theta_0 + \omega(a_0)t)$$

(periodic or quasi-periodic flow).

Linear error growth for integrable systems

Assumptions

- completely integrable Hamiltonian system with real-analytic Hamiltonian, action variables $a = I(p, q)$
- symplectic integrator of order r
- some technical assumptions.

Then, there exist constants C, h_0 such that for $h \le h_0$ and for $t = nh \le h^{-r}$ the numerical solution satisfies

$$\|(p_n, q_n) - (p(t), q(t))\| \le C\, t\, h^r \qquad \text{(linear error growth)}$$

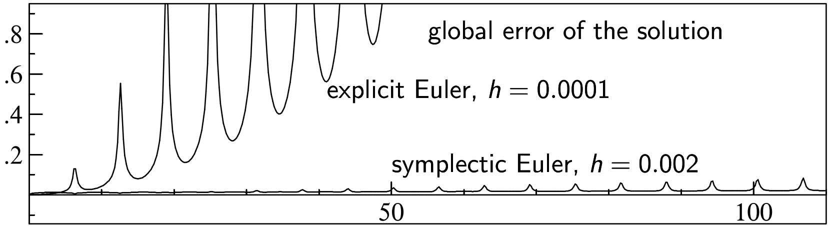
$$\|I(p_n, q_n) - I(p_0, q_0)\| \le C\, h^r \qquad \text{(near-conserv. of actions)}$$

Remark. The same statement is true if we replace
"Hamiltonian" by "reversible" and
"symplectic" by "symmetric".

Numerical experiment

Kepler problem (excentricity $e = 0.6$)

initial values, such that the orbit is elliptic with period 2π



Explicit Euler: quadratic error growth

Symplectic Euler: linear error growth

Lecture 2. Backward error analysis

Charged particle dynamics

Newton's Second Law together with Lorentz's force equation yields (assuming suitable units)

$$\ddot{x} = \dot{x} \times B(x) + E(x)$$

where $E(x)$ is the electric field and $B(x)$ the magnetic field.

Boris algorithm

The most simple discretization is

$$x_{n+1} - 2x_n + x_{n-1} = \frac{h}{2}(x_{n+1} - x_{n-1}) \times B(x_n) + h^2 E(x_n)$$

J.P. Boris, *Relativistic plasma simulation-optimization of a hybrid code.* Proc. of 4th Conf. on Numer. Simul. of Plasmas (Nov. 1970)

Properties of the differential equation

We write

$$\ddot{x} = \dot{x} \times B(x) + E(x)$$

as

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= v \times B(x) + E(x) \end{aligned}$$

- the flow $\varphi_t(x, v)$ is volume preserving:

$$\mu\big(\varphi_t(K)\big) = \mu(K) \qquad \text{for all } t;$$

- if $E(x) = -\nabla U(x)$, the energy

$$H(x, v) = \tfrac{1}{2} v^{\top} v + U(x) \qquad \text{is preserved;}$$

- if $E(x) = -\nabla U(x)$ and $B(x) = \nabla_x \times A(x)$, the differential equations are the Euler–Lagrange equations with

$$L(x, v) = \tfrac{1}{2} v^{\top} v - U(x) + A(x)^{\top} v$$

Properties of the differential equation

We write

$$\ddot{x} = \dot{x} \times B(x) + E(x)$$

as

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= v \times B(x) + E(x) \end{aligned}$$

- the flow $\varphi_t(x, v)$ is volume preserving:

$$\mu\big(\varphi_t(K)\big) = \mu(K) \qquad \text{for all } t;$$

- if

Proof. Divergence of the vector field $= 0$, because $v \times B(x) = \widehat{B}(x)\, v$ with a skew-symmetric matrix $\widehat{B}(x)$.

- if $E(x) = -\nabla U(x)$ and $B(x) = \nabla_x \times A(x)$, the differential equations are the Euler–Lagrange equations with

$$L(x, v) = \tfrac{1}{2} v^{\mathsf{T}} v - U(x) + A(x)^{\mathsf{T}} v$$

Properties of the differential equation

We write

$$\ddot{x} = \dot{x} \times B(x) + E(x)$$

as

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= v \times B(x) + E(x) \end{aligned}$$

- the flow $\varphi_t(x, v)$ is volume preserving:

$$\mu\big(\varphi_t(K)\big) = \mu(K) \qquad \text{for all } t;$$

- if $E(x) = -\nabla U(x)$, the energy

$$H(x, v) = \tfrac{1}{2} v^\top v + U(x) \qquad \text{is preserved;}$$

- if $E(x) = -\nabla U(x)$ and $B(x) = \nabla_x \times A(x)$, the differential equations are the Euler–Lagrange equations with

$$L(x, v) = \tfrac{1}{2} v^\mathsf{T} v - U(x) + A(x)^\mathsf{T} v$$

Properties of the differential equation

We write

$$\ddot{x} = \dot{x} \times B(x) + E(x)$$

as

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= v \times B(x) + E(x) \end{aligned}$$

- the flow $\varphi_t(x, v)$ is volume preserving:

$$\mu\big(\varphi_t(K)\big) = \mu(K) \qquad \text{for all } t;$$

- if $E(x) = -\nabla U(x)$, the energy

$$H(x, v) = \tfrac{1}{2} v^\top v + U(x) \qquad \text{is preserved;}$$

- if … ations
  ar…

Proof.

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t} H\big(x(t), v(t)\big) &= v^\top \dot{v} + \dot{x}^\top \nabla U(x) \\ &= v^\top \big(v \times B(x) - \nabla U(x)\big) + v^\top \nabla U(x) = 0 \end{aligned}$$

Properties of the differential equation

We write

$$\ddot{x} = \dot{x} \times B(x) + E(x)$$

as

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= v \times B(x) + E(x) \end{aligned}$$

- the flow $\varphi_t(x, v)$ is volume preserving:

$$\mu\big(\varphi_t(K)\big) = \mu(K) \qquad \text{for all } t;$$

- if $E(x) = -\nabla U(x)$, the energy

$$H(x, v) = \tfrac{1}{2} v^\top v + U(x) \qquad \text{is preserved;}$$

- if $E(x) = -\nabla U(x)$ and $B(x) = \nabla_x \times A(x)$, the differential equations are the Euler–Lagrange equations with

$$L(x, v) = \tfrac{1}{2} v^\mathsf{T} v - U(x) + A(x)^\mathsf{T} v$$

Properties of the differential equation

We write

$$\ddot{x} = \dot{x} \times B(x) + E(x)$$

as

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= v \times B(x) + E(x) \end{aligned}$$

- the flow $\varphi_t(x,v)$ is volume preserving;

- if

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(\nabla_v L\big) = \nabla_x L$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(v + A(x)\big) = -\nabla_x U + \nabla_x\big(A(x)^\top v\big)$$

and the statement follows from

$$\nabla_x\big(A(x)^\top v\big) - \frac{\mathrm{d}}{\mathrm{d}t}A(x) = \big(A'(x)^\top - A'(x)\big)v = v \times B(x)$$

- if $E(x) = -\nabla U(x)$ and $B(x) = \nabla_x \times A(x)$, the differential equations are the Euler–Lagrange equations with

$$L(x,v) = \frac{1}{2}v^\mathsf{T} v - U(x) + A(x)^\mathsf{T} v$$

Properties of the differential equation

We write

$$\ddot{x} = \dot{x} \times B(x) + E(x) \quad \text{as} \quad \begin{aligned} \dot{x} &= v \\ \dot{v} &= v \times B(x) + E(x) \end{aligned}$$

- the flow $\varphi_t(x, v)$ is volume preserving:

$$\mu\big(\varphi_t(K)\big) = \mu(K) \qquad \text{for all } t;$$

- if $E(x) = -\nabla U(x)$, the energy

$$H(x, v) = \tfrac{1}{2} v^\top v + U(x) \qquad \text{is preserved;}$$

- if $E(x) = -\nabla U(x)$ and $B(x) = \nabla_x \times A(x)$, the differential equations are the Euler–Lagrange equations with

$$L(x, v) = \tfrac{1}{2} v^\mathsf{T} v - U(x) + A(x)^\mathsf{T} v$$

Boris algorithm as one-step method

$$
\begin{aligned}
x_{n+1} - 2x_n + x_{n-1} &= \frac{h}{2}\,(x_{n+1} - x_{n-1}) \times B(x_n) + h^2 E(x_n) \\
v_n &= \frac{1}{2h}\,(x_{n+1} - x_{n-1})
\end{aligned}
$$

With $v_{n+1/2} = \frac{1}{h}(x_{n+1} - x_n) = v_n + \frac{h}{2}\, v_n \times B(x_n) + \frac{h}{2} E(x_n)$ we have

$$
v_{n+1/2} - v_{n-1/2} = \frac{h}{2}\,(v_{n+1/2} + v_{n-1/2}) \times B(x_n) + h\, E(x_n)
$$

and the map $(x_n, v_{n-1/2}) \mapsto (x_{n+1}, v_{n+1/2})$ is implemented as

$$
\begin{aligned}
v^+_{n-1/2} &= v_{n-1/2} + \frac{h}{2}\, E(x_n) \\
v^-_{n+1/2} - v^+_{n-1/2} &= \frac{h}{2}\,(v^-_{n+1/2} + v^+_{n-1/2}) \times B(x_n) \\
v_{n+1/2} &= v^-_{n+1/2} + \frac{h}{2}\, E(x_n) \\
x_{n+1} &= x_n + h\, v_{n+1/2}
\end{aligned}
$$

Boris algorithm as one-step method

$$\begin{aligned} x_{n+1} - 2x_n + x_{n-1} &= \frac{h}{2}\,(x_{n+1} - x_{n-1}) \times B(x_n) + h^2 E(x_n) \\ v_n &= \frac{1}{2h}\,(x_{n+1} - x_{n-1}) \end{aligned}$$

With $v_{n+1/2} = \frac{1}{h}(x_{n+1} - x_n) = v_n + \frac{h}{2}\,v_n \times B(x_n) + \frac{h}{2} E(x_n)$ we have

$$v_{n+1/2} - v_{n-1/2} = \frac{h}{2}\,(v_{n+1/2} + v_{n-1/2}) \times B(x_n) + h\,E(x_n)$$

and the map $(x_n, v_{n-1/2}) \mapsto (x_{n+1}, v_{n+1/2})$ is implemented as

$$\begin{aligned} v^+_{n-1/2} &= v_{n-1/2} + \frac{h}{2}\,E(x_n) \\ v^-_{n+1/2} - v^+_{n-1/2} &= \frac{h}{2}\,(v^-_{n+1/2} + v^+_{n-1/2}) \times B(x_n) \\ v_{n+1/2} &= v^-_{n+1/2} + \frac{h}{2}\,E(x_n) \\ x_{n+1} &= x_n + h\,v_{n+1/2} \end{aligned}$$

Boris algorithm as one-step method

$$\begin{aligned} x_{n+1} - 2x_n + x_{n-1} &= \tfrac{h}{2}\,(x_{n+1} - x_{n-1}) \times B(x_n) + h^2 E(x_n) \\ v_n &= \tfrac{1}{2h}\,(x_{n+1} - x_{n-1}) \end{aligned}$$

With $v_{n+1/2} = \frac{1}{h}(x_{n+1} - x_n) = v_n + \frac{h}{2}\, v_n \times B(x_n) + \frac{h}{2} E(x_n)$ we have

$$v_{n+1/2} - v_{n-1/2} = \tfrac{h}{2}\,(v_{n+1/2} + v_{n-1/2}) \times B(x_n) + h\, E(x_n)$$

and the map $(x_n, v_{n-1/2}) \mapsto (x_{n+1}, v_{n+1/2})$ is implemented as

$$\begin{aligned} v^+_{n-1/2} &= v_{n-1/2} + \tfrac{h}{2}\, E(x_n) \\ v^-_{n+1/2} - v^+_{n-1/2} &= \tfrac{h}{2}\,(v^-_{n+1/2} + v^+_{n-1/2}) \times B(x_n) \\ v_{n+1/2} &= v^-_{n+1/2} + \tfrac{h}{2}\, E(x_n) \\ x_{n+1} &= x_n + h\, v_{n+1/2} \end{aligned}$$

Boris algorithm as one-step method

With the splitting

$$\begin{pmatrix} \dot{x} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} 0 \\ E(x) \end{pmatrix} + \begin{pmatrix} 0 \\ v \times B(x) \end{pmatrix} + \begin{pmatrix} v \\ 0 \end{pmatrix}$$

we have

$$\begin{pmatrix} x_{n+1} \\ v_{n+1/2} \end{pmatrix} = \varphi_h^V \circ \varphi_{h/2}^E \circ \Phi_h^B \circ \varphi_{h/2}^E \begin{pmatrix} x_n \\ v_{n-1/2} \end{pmatrix}$$

where φ_t^E and φ_t^V are the exact flows, and Φ_h^B is the discrete flow (mid-point rule) for the vector field in the middle.

Wi

and

$$\begin{aligned} v_{n-1/2}^{+} &= v_{n-1/2} + \frac{h}{2} E(x_n) \\ v_{n+1/2}^{-} - v_{n-1/2}^{+} &= \frac{h}{2} \left(v_{n+1/2}^{-} + v_{n-1/2}^{+} \right) \times B(x_n) \\ v_{n+1/2} &= v_{n+1/2}^{-} + \frac{h}{2} E(x_n) \\ x_{n+1} &= x_n + h\, v_{n+1/2} \end{aligned}$$

Properties of the Boris algorithm

$$
\begin{aligned}
x_{n+1} - 2x_n + x_{n-1} &= \tfrac{h}{2}\,(x_{n+1} - x_{n-1}) \times B(x_n) + h^2 E(x_n) \\
v_n &= \tfrac{1}{2h}\,(x_{n+1} - x_{n-1})
\end{aligned}
$$

- the mapping $(x_n, v_{n-1/2} \mapsto (x_{n+1}, v_{n+1/2})$ is volume preserving. Hence, the Boris method $(x_n, v_n) \mapsto (x_{n+1}, v_{n+1})$ is conjugate to a volume preserving mapping.
- the Boris method is a variational integrator only if $B(x) = Const.$ *(see Ellison & al., and part II of the talk)*
- What can be said about near energy preservation in the general case, where $B(x)$ is not a constant vector field? *This is the topic of the present talk.*

C. L. Ellison, J. W. Burby, and H. Qin, *Comment on "Symplectic integration of magnetic systems": A proof that the Boris algorithm is not variational. J. Comput. Phys. 301* (2015), 489–493

Properties of the Boris algorithm

$$x_{n+1} - 2x_n + x_{n-1} = \frac{h}{2}\,(x_{n+1} - x_{n-1}) \times B(x_n) + h^2 E(x_n)$$
$$v_n = \frac{1}{2h}\,(x_{n+1} - x_{n-1})$$

- the mapping $(x_n, v_{n-1/2} \mapsto (x_{n+1}, v_{n+1/2})$ is volume preserving. Hence, the Boris method $(x_n, v_n) \mapsto (x_{n+1}, v_{n+1})$ is conjugate to a volume preserving mapping.
- the Boris method is a variational integrator only if $B(x) = Const$. *(see Ellison & al., and part II of the talk)*
- What can be said about near energy preservation in the general case, where $B(x)$ is not a constant vector field? *This is the topic of the present talk.*

C. L. Ellison, J. W. Burby, and H. Qin, *Comment on "Symplectic integration of magnetic systems": A proof that the Boris algorithm is not variational. J. Comput. Phys. 301* (2015), 489–493

Properties of the Boris algorithm

$$x_{n+1} - 2x_n + x_{n-1} = \frac{h}{2}\left(x_{n+1} - x_{n-1}\right) \times B(x_n) + h^2 E(x_n)$$
$$v_n = \frac{1}{2h}\left(x_{n+1} - x_{n-1}\right)$$

- the mapping $(x_n, v_{n-1/2} \mapsto (x_{n+1}, v_{n+1/2})$ is volume preserving. Hence, the Boris method $(x_n, v_n) \mapsto (x_{n+1}, v_{n+1})$ is conjugate to a volume preserving mapping.
- the Boris method is a variational integrator only if $B(x) = Const$. *(see Ellison & al., and part II of the talk)*
- What can be said about near energy preservation in the general case, where $B(x)$ is not a constant vector field?
  This is the topic of the present talk.

C. L. Ellison, J. W. Burby, and H. Qin, *Comment on "Symplectic integration of magnetic systems": A proof that the Boris algorithm is not variational. J. Comput. Phys. 301* (2015), 489–493

Energy preservation - main result

Theorem

Assume that at least one of the following conditions is satisfied

- *the magnetic field* $B(x) = B$ *is constant,*
- *the scalar potential* $U(x) = \frac{1}{2} x^\top Q x + q^\top x$ *is quadratic,*

and that the numerical solution (x_n, v_n) *of the Boris method stays in a compact set. For every truncation index* N*, the energy* $H(x,v) = \frac{1}{2} v^\top v + U(x)$ *is bounded as*

$$|H(x_n, v_n) - H(x_0, v_0)| \le C_{2N} h^2 \quad \text{for} \quad nh \le h^{-2N}$$

with C *independent of* n *and* h *as long as* $nh \le h^{-2N}$.

What happens if none of the above two conditions is satisfied?

E. Hairer and Ch. Lubich, *Energy behaviour of the Boris method for charged-particle dynamics. BIT* (2018)

Energy preservation - main result

Theorem

Assume that at least one of the following conditions is satisfied

- *the magnetic field* $B(x) = B$ *is constant,*
- *the scalar potential* $U(x) = \frac{1}{2} x^\top Q x + q^\top x$ *is quadratic,*

and that the numerical solution (x_n, v_n) *of the Boris method stays in a compact set. For every truncation index* N*, the energy* $H(x, v) = \frac{1}{2} v^\top v + U(x)$ *is bounded as*

$$|H(x_n, v_n) - H(x_0, v_0)| \leq C_{2N} h^2 \quad \text{for} \quad nh \leq h^{-2N}$$

with C *independent of* n *and* h *as long as* $nh \leq h^{-2N}$*.*

What happens if none of the above two conditions is satisfied?

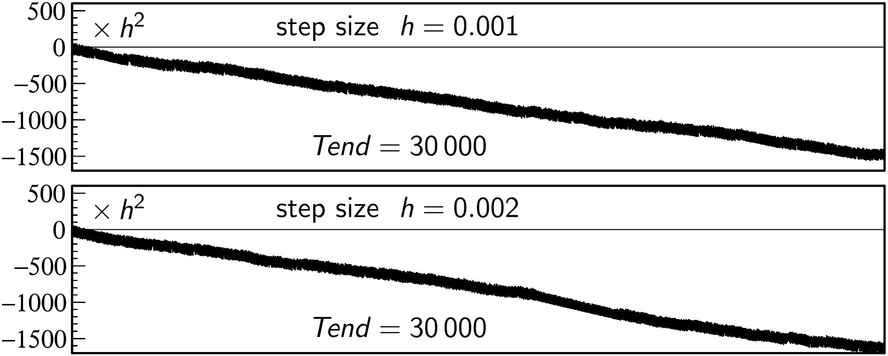
E. Hairer and Ch. Lubich, *Energy behaviour of the Boris method for charged-particle dynamics. BIT* (2018)

Example 1: linear growth

We consider the error in the energy for

$$U(x) = x_1^3 - x_2^3 + \frac{1}{5}x_1^4 + x_2^4 + x_3^4, \qquad B(x) = \begin{pmatrix} 0 \\ 0 \\ \sqrt{x_1^2 + x_2^2} \end{pmatrix}$$

$x(0) = (0.0, 1.0, 0.1)^\top$, $v(0) = (0.09, 0.55.0.30)^\top$.

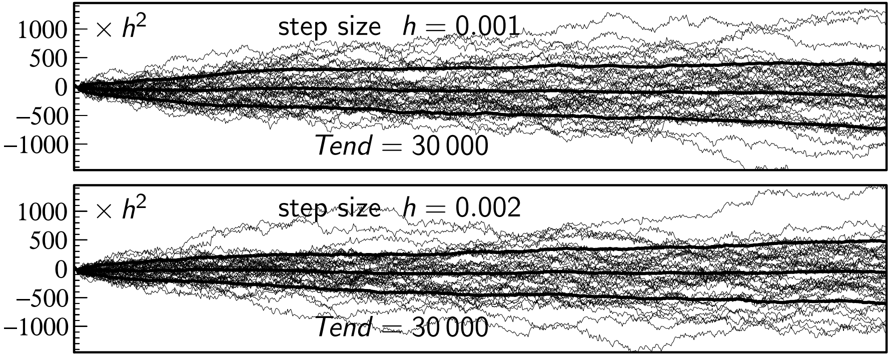


Example 2: random walk

We consider the error in the energy for

$$U(x) = x_1^3 - x_2^3 + \frac{1}{5}x_1^4 + x_2^4 + x_3^4, \qquad B(x) = \frac{1}{2}\begin{pmatrix} x_2 - x_3 \\ x_1 + x_3 \\ x_2 - x_1 \end{pmatrix}$$

$x(0) = (0.0, 1.0, 0.1)^\top$, $v(0) = (0.09, 0.55.0.30)^\top$.



Backward error analysis (Boris algorithm)

For $x_n = y(nh)$ and $t = nh$ the Boris algorithm reads

$$y(t+h)-2y(t)+y(t-h) = \frac{h}{2}\Big(y(t+h)-y(t-h)\Big)\times B\big(y(t)\big)-h^2\nabla U\big(y(t)\big)$$

Expanding into powers of h and dividing by h^2 yields

$$\ddot{y} + \frac{h^2}{12}\ddddot{y} + \ldots = \Big(\dot{y} + \frac{h^2}{6}\dddot{y} + \ldots\Big)\times B(y) - \nabla U(y)$$

Eliminating third and higher derivatives by differentiation

$$\begin{aligned} \dddot{y} &= \ddot{y}\times B(y) + \dot{y}\times B'(y)\dot{y} - \nabla^2 U(y)\dot{y} + \mathcal{O}(h^2) \\ &= -\nabla U(y)\times B(y) + \dot{y}\times B'(y)\dot{y} - \nabla^2 U(y)\dot{y} + \mathcal{O}(h^2) \end{aligned}$$

gives the modified differential equation.

Similarly, we have $v_n = w(nh)$ for $t = nh$, where

$$w(t) = \frac{1}{2h}\Big(y(t+h)-y(t-h)\Big) = \dot{y} + \frac{h^2}{3!}\dddot{y} + \frac{h^4}{5!}y^{(5)} + \ldots$$

Backward error analysis (Boris algorithm)

For $x_n = y(nh)$ and $t = nh$ the Boris algorithm reads

$$y(t+h)-2y(t)+y(t-h) = \frac{h}{2}\Big(y(t+h)-y(t-h)\Big)\times B\big(y(t)\big)-h^2\nabla U\big(y(t)\big)$$

Expanding into powers of h and dividing by h^2 yields

$$\ddot y + \frac{h^2}{12}\ddddot y + \ldots = \Big(\dot y + \frac{h^2}{6}\dddot y + \ldots\Big)\times B(y) - \nabla U(y)$$

Eliminating third and higher derivatives by differentiation

$$\begin{aligned}\dddot y &= \ddot y \times B(y) + \dot y \times B'(y)\dot y - \nabla^2 U(y)\dot y + \mathcal{O}(h^2)\\ &= -\nabla U(y)\times B(y) + \dot y \times B'(y)\dot y - \nabla^2 U(y)\dot y + \mathcal{O}(h^2)\end{aligned}$$

gives the modified differential equation.

Similarly, we have $v_n = w(nh)$ for $t = nh$, where

$$w(t) = \frac{1}{2h}\Big(y(t+h)-y(t-h)\Big) = \dot y + \frac{h^2}{3!}\dddot y + \frac{h^4}{5!}y^{(5)} + \ldots$$

Energy conservation - constant magnetic field

Consider the modified equation

$$\ddot{y} + \frac{h^2}{12}\ddddot{y} + \ldots = \Big(\dot{y} + \frac{h^2}{6}\dddot{y} + \ldots\Big) \times B(y) - \nabla U(y)$$

and take the scalar product with $\dot{y}$. This gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}\dot{y}^\top\dot{y} + U(y) + \frac{h^2}{12}\big(\dot{y}^\top\dddot{y} - \frac{1}{2}\ddot{y}^\top\ddot{y}\big) + \ldots\Big) = \frac{h^2}{6}\dot{y}^\top\big(\dddot{y} \times B(y)\big) + \ldots$$

Theorem

If the magnetic field $B(x) = B$ is constant, there exist $E_{2j}(x, v)$ such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}\dot{y}^\top\dot{y} + U(y) + h^2 E_2(y,\dot{y}) + \ldots + h^{2N} E_{2N}(y,\dot{y})\Big) = \mathcal{O}(h^{2N+2})$$

along solutions $(y, \dot{y})$ of the modified differential equation.

Energy conservation - constant magnetic field

Consider the modified equation

$$\ddot{y} + \frac{h^2}{12}\ddddot{y} + \ldots = \Big(\dot{y} + \frac{h^2}{6}\dddot{y} + \ldots\Big) \times B(y) - \nabla U(y)$$

and take the scalar product with $\dot{y}$. This gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}\dot{y}^\top\dot{y} + U(y) + \frac{h^2}{12}\big(\dot{y}^\top\dddot{y} - \frac{1}{2}\ddot{y}^\top\ddot{y}\big) + \ldots\Big) = \frac{h^2}{6}\dot{y}^\top\big(\dddot{y} \times B(y)\big) + \ldots$$

Theorem

If the magnetic field $B(x) = B$ is constant, there exist $E_{2j}(x,v)$ such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}\dot{y}^\top\dot{y} + U(y) + h^2 E_2(y,\dot{y}) + \ldots + h^{2N}E_{2N}(y,\dot{y})\Big) = \mathcal{O}(h^{2N+2})$$

along solutions $(y,\dot{y})$ of the modified differential equation.

Energy conservation - constant magnetic field

Consider the modified equation

$$\ddot y + \frac{h^2}{12}\, y^{(4)} + \ldots = \Big(\dot y + \frac{h^2}{6}\,\dddot y + \ldots\Big) \times B(y) - \nabla U(y)$$

and ta

because $\quad \dot y^\top(\dddot y \times B(y)) = \dot y^\top(\dddot y \times B) = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\dot y^\top(\ddot y \times B)\Big)$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}\dot y^\top \dot y + U(y) + \frac{h^2}{12}\big(\dot y^\top \dddot y - \frac{1}{2}\ddot y^\top \ddot y\big) + \ldots\Big) = \frac{h^2}{6}\dot y^\top(\dddot y \times B(y)) + \ldots$$

Theorem

If the magnetic field $B(x) = B$ is constant, there exist $E_{2j}(x, v)$ such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}\dot y^\top \dot y + U(y) + h^2 E_2(y, \dot y) + \ldots + h^{2N} E_{2N}(y, \dot y)\Big) = \mathcal{O}(h^{2N+2})$$

along solutions $(y, \dot y)$ of the modified differential equation.

Energy conservation - quadratic electric potential

Consider the modified equation with $U(x) = \frac{1}{2}\, x^\top Q x + q^\top x$

$$\ddot{y} + \frac{h^2}{12}\ddddot{y} + \ldots = \Big(\dot{y} + \frac{h^2}{6}\dddot{y} + \ldots\Big) \times B(y) - \nabla U(y)$$

and take the scalar product with $\big(\dot{y} + \frac{h^2}{6}\dddot{y} + \ldots\big)$. This gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}\dot{y}^\top\dot{y} + U(y) + \frac{h^2}{12}\big(\dot{y}^\top\dddot{y} + \frac{1}{2}\ddot{y}^\top\ddot{y}\big) + \ldots\Big) = -\frac{h^2}{6}\dddot{y}^\top\nabla U(y) + \ldots$$

Theorem

If $U(x) = \frac{1}{2}\, x^\top Q x + q^\top x$, *there exist* $E_{2j}(x,v)$ *such that*

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}\dot{y}^\top\dot{y} + U(y) + h^2 E_2(y,\dot{y}) + \ldots + h^{2N} E_{2N}(y,\dot{y})\Big) = \mathcal{O}(h^{2N+2})$$

along solutions (y,w) *of the modified differential equation.*

Energy conservation - quadratic electric potential

Consider the modified equation with $U(x) = \frac{1}{2}\, x^\top Q x + q^\top x$

$$\ddot y + \frac{h^2}{12}\, \ddddot y + \ldots = \Big(\dot y + \frac{h^2}{6}\, \dddot y + \ldots \Big) \times B(y) - \nabla U(y)$$

and take the scalar product with $\big(\dot y + \frac{h^2}{6}\, \dddot y + \ldots \big)$. This gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2} \dot y^\top \dot y + U(y) + \frac{h^2}{12} \big(\dot y^\top \dddot y + \frac{1}{2} \ddot y^\top \ddot y \big) + \ldots \Big) = -\frac{h^2}{6}\, \dddot y^\top \nabla U(y) + \ldots$$

Theorem

If $U(x) = \frac{1}{2}\, x^\top Q x + q^\top x$, *there exist* $E_{2j}(x, v)$ *such that*

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2} \dot y^\top \dot y + U(y) + h^2 E_2(y, \dot y) + \ldots + h^{2N} E_{2N}(y, \dot y) \Big) = \mathcal{O}(h^{2N+2})$$

along solutions (y, w) *of the modified differential equation.*

Energy conservation - quadratic electric potential

Consider the modified equation with $U(x) = \frac{1}{2}\, x^\top Q x + q^\top x$

$$\ddot y + \frac{h^2}{12}\, y^{(4)} + \ldots = \Big(\dot y + \frac{h^2}{6}\, \dddot y + \ldots \Big) \times B(y) - \nabla U(y)$$

because $\quad \dddot y^\top \nabla U(y) = \dddot y^\top (Qy + q) = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\ddot y^\top Q y - \frac{1}{2} \dot y^\top Q \dot y + \ddot y^\top q \Big)$

ar

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2} \dot y^\top \dot y + U(y) + \frac{h^2}{12}\big(\dot y^\top \dddot y + \frac{1}{2} \ddot y^\top \ddot y \big) + \ldots \Big) = -\frac{h^2}{6}\, \dddot y^\top \nabla U(y) + \ldots$$

Theorem

If $U(x) = \frac{1}{2}\, x^\top Q x + q^\top x$, *there exist* $E_{2j}(x, v)$ *such that*

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2} \dot y^\top \dot y + U(y) + h^2 E_2(y, \dot y) + \ldots + h^{2N} E_{2N}(y, \dot y) \Big) = \mathcal{O}(h^{2N+2})$$

along solutions (y, w) *of the modified differential equation.*

Explanation of the numerical experiments

The numerical solution satisfies (formally) $x_n = y(nh)$, $v_n = w(nh)$, where

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(H(y,w)+h^2F(y,w)+\ldots\Big) = h^2G(y,w)+\mathcal{O}(h^4)$$

- integration shows that there is typically a linear drift of size th^2M, where M is an upper bound of $G(y,w)$,
- if the solution $(y(t),w(t))$ is ergodic on an invariant set A with invariant measure μ, then

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t G\big(y(s),w(s)\big)\,\mathrm{d}s = \int_A G(x,v)\mu\big(d(x,v)\big)$$

  - if the integral to the right is non-zero, we will have a linear drift;
  - if it is zero, we will have a random walk: the error behaves like $\mathcal{O}(h^2)+\mathcal{O}(\sqrt{t}\,h^2)$.