Structure-Preserving Discretizations of Initial-Value Problems

John C. Bowman, Department of Mathematical and Statistical Sciences,
University of Alberta, Edmonton, Alberta, Canada T6G 2G1
http://www.math.ualberta.ca/∼bowman/talks

Outline

Initial Value Problems

- Given $\boldsymbol{f} : \mathbb{R}^{n+1} \to \mathbb{R}^n$, suppose $\boldsymbol{x} \in \mathbb{R}^n$ evolves according to

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(\boldsymbol{x}, t), \tag{1}$$

  with the initial condition $\boldsymbol{x}(0) = \boldsymbol{x}_0$.

- If $n = 2k$ and $\boldsymbol{x} = (\boldsymbol{p}, \boldsymbol{q})$ where $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{R}^k$ satisfy

$$\frac{d\boldsymbol{q}}{dt} = \frac{\partial H}{\partial \boldsymbol{p}},$$

$$\frac{d\boldsymbol{p}}{dt} = -\frac{\partial H}{\partial \boldsymbol{q}},$$

  for some function $H(\boldsymbol{p}, \boldsymbol{q}, t) : \mathbb{R}^{n+1} \to \mathbb{R}$, we say that (1) is Hamiltonian.

- Often, the Hamiltonian H has no explicit dependence on t.

Structure-Preserving Discretizations

- Symplectic integration: conserves phase space structure of Hamilton's equations; the time step map is a canonical transformation. [Ruth 1983, Channell & Scovel 1990, Sanz-Serna & Calvo 1994]
- Conservative integration: conserves first integrals. [Bowman *et al.* 1997, Shadwick *et al.* 1999, Kotovych & Bowman 2002]
- Positivity: preserves positive semi-definiteness of covariance matrices. [Bowman *et al.* 1993, Bowman & Krommes 1997]
- Unitary integration: conserves trace of probability density matrix. [Shadwick & Buell 1997]
- Operator splitting: e.g. to yield exact evolution on linear time scale.

Symplectic vs. Conservative Integration

Theorem 1 (Ge and Marsden 1988): *A C^1 symplectic map M with no explicit time-dependence will conserve a C^1 time-independent Hamiltonian $H : \mathbb{R}^n \to \mathbb{R} \iff M$ is identical to the exact evolution, up to a reparametrization of time.*

Proof:

- A C^1 symplectic scheme is a canonical map M corresponding to some approximate C^1 Hamiltonian $\tilde{H}_\tau(\boldsymbol{x}, t) : \mathbb{R}^{n+1} \to \mathbb{R}$, where the label τ denotes the time step.
- If the mapping M does not depend explicitly on time, it can be generated by the approximate Hamiltonian $K(\boldsymbol{x}) = \tilde{H}_\tau(\boldsymbol{x}, 0)$.

- Suppose the symplectic map conserves the true Hamiltonian H:

$$0 = \frac{dH}{dt} = \frac{\partial H}{\partial q_i}\frac{dq_i}{dt} + \frac{\partial H}{\partial p_i}\frac{dp_i}{dt} + \cancel{\frac{\partial H}{\partial t}} = [H, K],$$

  where

$$[H, K] = \frac{\partial H}{\partial q_i}\frac{\partial K}{\partial p_i} - \frac{\partial H}{\partial p_i}\frac{\partial K}{\partial q_i}.$$

- Implicit function theorem: in a neighbourhood of $\boldsymbol{x}_0 \in \mathbb{R}^n$ $\exists$ a C^1 function $\phi : \mathbb{R} \to \mathbb{R} \ni$

$$H(\boldsymbol{x}) = \phi(K(\boldsymbol{x})) \quad \text{or} \quad K(\boldsymbol{x}) = \phi(H(\boldsymbol{x})) \iff [H, K] = 0.$$

- Consequently, the trajectories in $\mathbb{R}^n$ generated by the Hamiltonians H and K coincide.

$Q.E.D.$

Conservative Integration

- Traditional numerical discretizations of nonlinear initial value problems, based on polynomial functions of the time step, typically yield spurious secular drifts of nonlinear first integrals of motion (such as the total energy).
  $\Rightarrow$ the numerical solution will not remain on the energy surface defined by the initial conditions!
- There exists a class of nontraditional explicit algorithms that exactly conserve nonlinear invariants to $all\ orders$ in the time step (to machine precision).

Three-Wave Problem

- Truncated Fourier-transformed Euler equations for an inviscid 2D fluid:

$$\frac{dx_1}{dt} = f_1 = M_1 x_2 x_3,$$
$$\frac{dx_2}{dt} = f_2 = M_2 x_3 x_1,$$
$$\frac{dx_3}{dt} = f_3 = M_3 x_1 x_2,$$

  where $M_1 + M_2 + M_3 = 0.$

- Then

$$\sum_k f_k x_k = 0 \Rightarrow \text{ energy } E \doteq \frac{1}{2} \sum_k x_k^2 \text{ is conserved.}$$

Secular Energy Growth

- Energy is not conserved by conventional discretizations like Euler, Predictor–Corrector, Runge–Kutta, ...
- The Euler method,

$$x_k(t+\tau) = x_k(t) + \tau f_k,$$

yields a monotonically increasing new energy:

$$E(t+\tau) = \frac{1}{2}\sum_k \left[x_k^2 + 2\tau f_k x_k + \tau^2 S_k^2\right]$$

$$= E(t) + \frac{1}{2}\tau^2 \sum_k S_k^2.$$

Conservative Euler Algorithm

- Try to determine a modification of the original equations of motion that will lead to $exact$ energy conservation:

$$\frac{dx_k}{dt} = f_k + g_k.$$

- Euler's method predicts the new energy

$$E(t+\tau) = \frac{1}{2}\sum_k \left[x_k + \tau(f_k + g_k)\right]^2$$

$$= E(t) + \frac{1}{2}\sum_k \underbrace{\left[2\tau g_k x_k + \tau^2 (f_k + g_k)^2\right]}_{\text{set to } 0}.$$

- Solving for g_k yields the C–Euler discretization:

$$x_k(t+\tau) = \operatorname{sgn} x_k(t+\tau)\,\sqrt{x_k^2 + 2\tau f_k x_k},$$

  which conserves energy exactly.

- As $\tau \to 0$, this reduces to Euler's method:

$$\begin{aligned} x_k(t+\tau) &= x_k\sqrt{1 + 2\tau \frac{f_k}{x_k}} \\ &= x_k + \tau f_k + \mathcal{O}(\tau^2). \end{aligned}$$

- C–Euler is just the usual Euler algorithm applied to

$$\frac{dx_k^2}{dt} = 2 f_k x_k.$$

Lemma 1: *Let* $\boldsymbol{x}$ *and* $\boldsymbol{c}$ *be vectors in* $\mathbb{R}^n$. *If* $\boldsymbol{f} : \mathbb{R}^{n+1} \to \mathbb{R}^n$ *has values orthogonal to* $\boldsymbol{c}$, *so that* $I = \boldsymbol{c} \cdot \boldsymbol{x}$ *is a linear invariant of*

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(\boldsymbol{x}, t),$$

then each stage of the explicit m*-stage discretization*

$$\boldsymbol{x}_j = \boldsymbol{x}_0 + \tau \sum_{k=0}^{j-1} b_{jk} \boldsymbol{f}(\boldsymbol{x}_k, t + a_j \tau), \qquad j = 1, \ldots, m,$$

also conserves I, *where* τ *is the time step and* $b_{jk} \in \mathbb{R}$.

Proof. For $j = 1, \ldots, m$, we have

$$\boldsymbol{c} \cdot \boldsymbol{x}_j = \boldsymbol{c} \cdot \boldsymbol{x}_0 + \tau \sum_{k=0}^{j-1} b_{jk} \boldsymbol{c} \cdot \boldsymbol{f}(\boldsymbol{x}_k, t + a_j \tau) = \boldsymbol{c} \cdot \boldsymbol{x}_0.$$

Predictor–Corrector (PC) Algorithm

- This second-order predictor–corrector (2-stage) scheme:

$$\tilde{\boldsymbol{x}} = \boldsymbol{x}_0 + \tau \boldsymbol{f}(\boldsymbol{x}_0, t),$$

$$\boldsymbol{x}(t+\tau) = \boldsymbol{x}_0 + \frac{\tau}{2}[\boldsymbol{f}(\boldsymbol{x}_0, t) + \boldsymbol{f}(\tilde{\boldsymbol{x}}, t+\tau)],$$

  conserves any invariant I that is a linear function of $\boldsymbol{x}$.

- Integration algorithms that conserve nonlinear invariants may be constructed by finding a transformation $\boldsymbol{T} : \mathbb{R}^n \to \mathbb{R}^n$ such that the nonlinear invariants are linear functions of $\boldsymbol{\xi} = \boldsymbol{T}(\boldsymbol{x})$.

- Retaining the original predictor

$$\tilde{\boldsymbol{x}} = \boldsymbol{x}_0 + \tau \boldsymbol{f}(\boldsymbol{x}_0, t),$$

  one computes the corrector in the transformed space,

$$\boldsymbol{\xi}(t+\tau) = \boldsymbol{\xi}_0 + \frac{\tau}{2}[\boldsymbol{T}'(\boldsymbol{x_0})\boldsymbol{f}(\boldsymbol{x}_0, t) + \boldsymbol{T}'(\tilde{\boldsymbol{x}})\boldsymbol{f}(\tilde{\boldsymbol{x}}, t+\tau)].$$

  where $\boldsymbol{T}'$ denotes the derivative of $\boldsymbol{T}$.

Conservative Predictor–Corrector (C–PC) Algorithm

- The new value of $\boldsymbol{x}$ is then obtained by inverse transformation:

$$\boldsymbol{x}(t+\tau) = \boldsymbol{T}^{-1}(\boldsymbol{\xi}(t+\tau)).$$

- Problem: T may not be invertible!
  Solution 1: Reduce the time step.
  Solution 2: Switch to a traditional integrator for that time step.
  Solution 3: Use an implicit backwards step [Shadwick & Bowman SIAM J. Appl. Math. **59**, 1112 (1999), Appendix A].

- Higher-order conservative integration algorithms are obtained by doing the final corrector stage in the transformed space:

$$\boldsymbol{\xi}(t+\tau) = \boldsymbol{\xi}_0 + \tau \sum_{k=0}^{m-1} b_{mk} \boldsymbol{T}'(\boldsymbol{x_k}) \boldsymbol{f}(\boldsymbol{x}_k, t + a_j\tau).$$

Error Analysis: 1D Autonomous Case

- Exact solution (everything on RHS evaluated at x_0):

$$x(t+\tau) = x_0 + \tau f + \frac{\tau^2}{2} f' f + \frac{\tau^3}{6}(f'' f^2 + f'^2 f) + \mathcal{O}(\tau^4);$$

- When $T'(x_0) \neq 0$, C–PC yields the solution

$$x(t+\tau) = x_0 + \tau f + \frac{\tau^2}{2} f' f + \frac{\tau^3}{4}\left(f'' f^2 + \frac{T'''}{3T'} f^3\right) + \mathcal{O}(\tau^4),$$

  where all of the derivatives are evaluated at x_0.

- On setting $T(x) = x$, the C–PC solution reduces to the conventional PC.

- C–PC and PC are both accurate to second order in τ; for $T(x) = x^2$, they agree through third order in τ.

Singular Case

- When $T'(x_0) = 0$, the conservative corrector reduces to

$$x(t+\tau) = T^{-1}\left(T(x_0) + \frac{\tau}{2}T'(\tilde{x})f(\tilde{x})\right),$$

- If T and f are analytic, the existence of a solution is guaranteed for sufficiently small positive τ, provided the points at which T' vanishes are isolated.

Example: Gravitational n-Body Problem

- Mass m_i is located at $\boldsymbol{r}_i$, $i = 1, \ldots, n$.
- Let $\boldsymbol{C}_i$ be the center of mass of the first i bodies.
- Enforce center of mass and linear momentum constraints: use Jacobi coordinates to obtain a reduced system of $n-1$ bodies at

$$\boldsymbol{\rho}_i = \boldsymbol{r}_i - \boldsymbol{C}_{i-1}, \qquad i = 2, \ldots, n,$$

  with center of mass at the origin.
- Let $M_j = \sum_{k=1}^{j-1} m_k$ and define the reduced masses

$$g_i = \frac{m_i M_{i-1}}{M_i}, \qquad i = 2, \ldots, n.$$

Hamiltonian Formulation

- The Hamiltonian is

$$H = \frac{1}{2} \sum_{i=2}^{n} \left(\frac{p_i^2}{g_i} + \frac{\ell_i^2}{g_i \rho_i^2} \right) + V,$$

where p_i and ℓ_i are the linear and angular momentum of the ith reduced mass and

$$V = - \sum_{\substack{i,j=1 \\ i<j}}^{n} \frac{G m_i m_j}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}.$$

Equations of Motion

- Both the energy H and the total angular momentum $L = \sum_{i=2}^{n} \ell_i$ are conserved by Hamilton's equations:

$$\dot{\rho}_i = \frac{\partial H}{\partial p_i} = \frac{p_i}{g_i},$$

$$\dot{\theta}_i = \frac{\partial H}{\partial \ell_i} = \frac{\ell_i}{g_i \rho_i^2},$$

$$\dot{p}_i = -\frac{\partial H}{\partial \rho_i} = \frac{\ell_i^2}{g_i \rho_i^3} - \frac{\partial V}{\partial \rho_i},$$

$$\dot{\ell}_i = -\frac{\partial H}{\partial \theta_i} = -\frac{\partial V}{\partial \theta_i},$$

where $i = 2, \ldots, n$ and the dots denote time derivatives.

Transformation

- We choose $\boldsymbol{T}$ to be the transformation [Kotovych & Bowman 2002]:

$$
\begin{aligned}
\zeta_2 &= V, \\
\zeta_i &= \rho_i, & i &= 3, \ldots, n, \\
\eta_i &= \frac{p_i^2}{2g_i} + \frac{\ell_i^2}{2g_i\rho_i^2}, & i &= 2, \ldots, n.
\end{aligned}
$$

- $H = \sum_{i=2}^{n} \eta_i + \zeta_2$ and $L = \sum_{i=2}^{n} \ell_i$ are linear functions of the transformed variables.

Corrector Equations

- The 2nd-order corrector equations are given by

$$\zeta_i(t+\tau) = \zeta_i + \frac{\tau}{2}(\dot{\zeta}_i + \dot{\tilde{\zeta}}_i), \qquad \theta_i(t+\tau) = \theta_i + \frac{\tau}{2}(\dot{\theta}_i + \dot{\tilde{\theta}}_i),$$

$$\eta_i(t+\tau) = \eta_i + \frac{\tau}{2}(\dot{\eta}_i + \dot{\tilde{\eta}}_i), \qquad \ell_i(t+\tau) = \ell_i + \frac{\tau}{2}(\dot{\ell}_i + \dot{\tilde{\ell}}_i),$$

where

$$\dot{\zeta}_2 = \sum_{i=2}^{n} \left(\frac{\partial V}{\partial \rho_i}\dot{\rho}_i + \frac{\partial V}{\partial \theta_i}\dot{\theta}_i \right),$$

$$\dot{\zeta}_i = \dot{\rho}_i, \qquad i = 3, \ldots, n,$$

$$\dot{\eta}_i = \frac{p_i \dot{p}_i}{g_i} + \frac{\ell_i \rho_i^2 \dot{\ell}_i - \rho_i \ell_i^2 \dot{\rho}_i}{g_i \rho_i^4}, \qquad i = 2, \ldots, n.$$

- One then inverts to get the original variables as functions of the temporary transformed variables:

$$\rho_i = \zeta_i, \qquad\qquad i = 3, \ldots, n,$$

$$\rho_2 = g(\zeta_2, \rho_3, \ldots, \rho_n, \boldsymbol{\theta}),$$

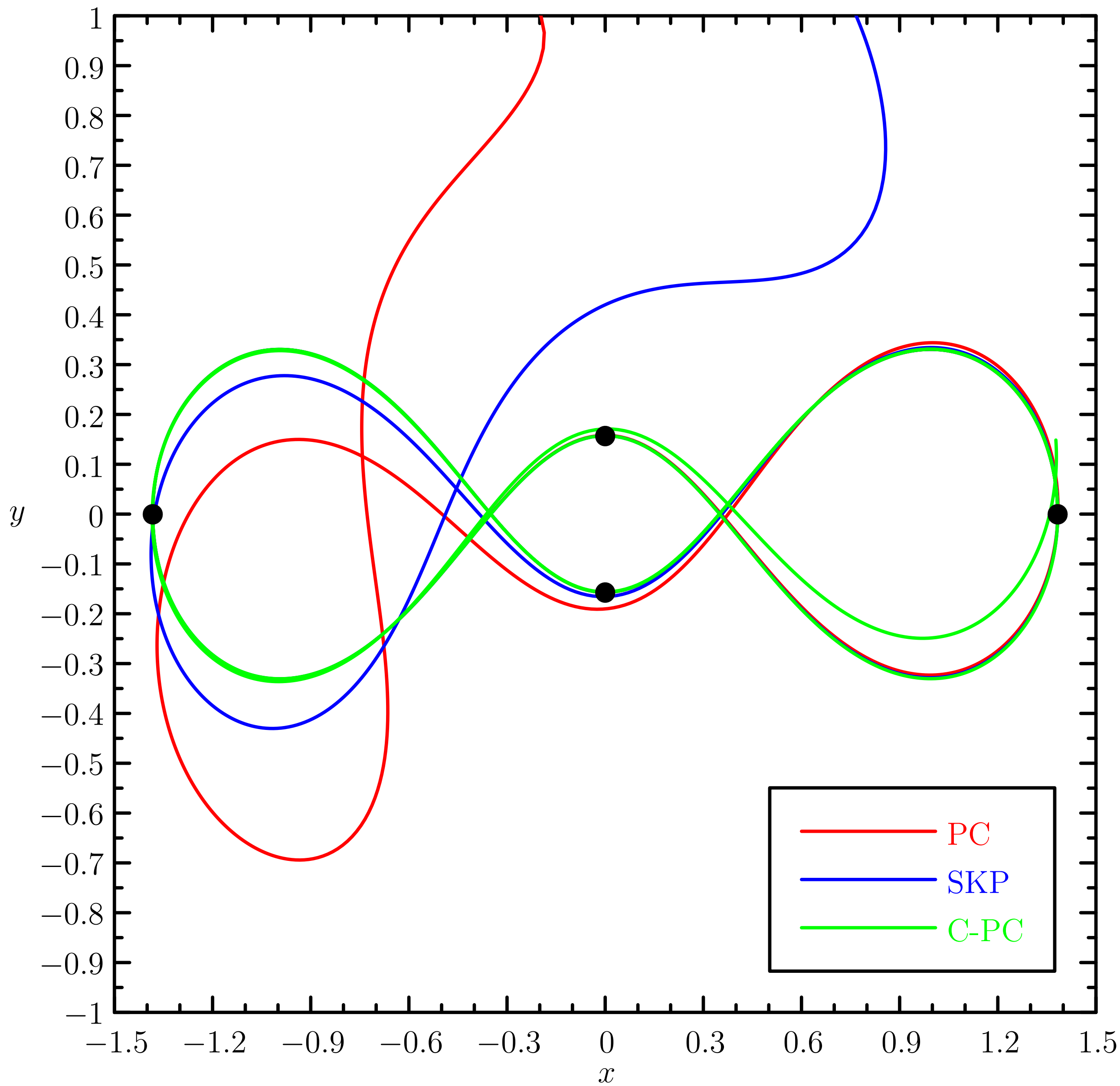
$$p_i = \operatorname{sgn}(\tilde{p}_i)\sqrt{2g_i\left(\eta_i - \frac{\ell_i^2}{2g_i\rho_i^2}\right)}, \qquad i = 2, \ldots, n.$$

- The value of the inverse function g defined by

$$V(g(\zeta_2, \rho_3, \ldots, \rho_n, \boldsymbol{\theta}), \rho_3, \ldots, \rho_n, \boldsymbol{\theta}) = \zeta_2$$

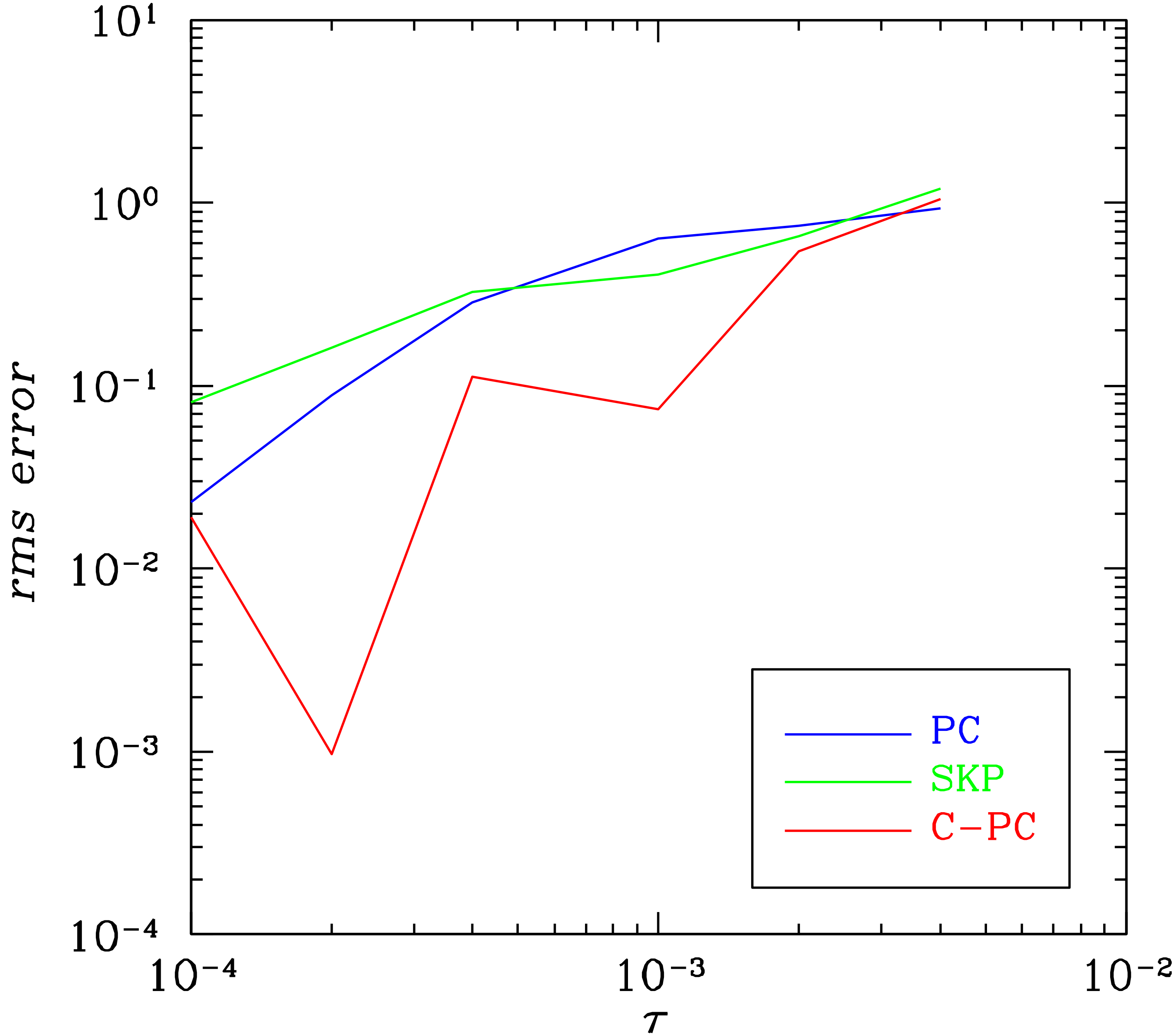
  is determined at fixed $\rho_3, \ldots, \rho_n$, $\boldsymbol{\theta}$ with a Newton–Raphson method, using the predicted value $\tilde{\rho}_2$ as an initial guess.

Four-body choreography



PC, symplectic SKP, and C–PC solutions

RMS error



PC, symplectic SKP, and C–PC errors

Conservative Symplectic Integrators

- Conservative variational symplectic integrators based on explicitly time-dependent symplectic maps have recently been developed for certain problems in mechanics.
- This allows one to circumvent the conditions of the Ge–Marsden theorem [Kane, Marsden, and Ortiz 1999].

Operator Splitting

- Typical stiff nonlinear initial value problem:

$$\frac{\partial x}{\partial t} + \eta\, x = S(t, x), \qquad x(0) = x_0.$$

- Stiff: Nonlinearity S has a slow variation in t compared with the value of the linear coefficient η:

$$\left| \frac{1}{S} \frac{dS}{dt} \right| \ll |\eta| \,.$$

- Goal: Solve on the linear time scale exactly; avoid the linear time-step restriction $\eta\tau \ll 1$.

- In the presence of nonlinearity, straightforward integrating factor methods do not remove the explicit restriction on the linear time step τ.

Exponential Euler Algorithm

- Exact evolution of x:

$$x(t_0+\tau) = P^{-1}(t_0+\tau)\left[x(t_0) + \int_{t_0}^{t_0+\tau} dt\, P(t)S(t)\right],$$

where $P(t) = e^{\eta(t-t_0)}$.

- Change variables: $dt\, P = \eta_0^{-1} dP \Rightarrow$

$$x(t_0+\tau) = P^{-1}(t_0+\tau)\left[x(t_0) + \eta_0^{-1}\int_1^{P(t_0+\tau)} dP\, S\right].$$

Rectangular approximation of integral $\Rightarrow$ Exponential Euler algorithm:

$$x_{i+1} = P_{i+1}^{-1}\left[x_i + \eta_0^{-1}(P_{i+1}-1)S_i\right].$$

- The discretization is now with respect to P instead of t.
- Also known as the Exponentially Fitted Euler method.

Generalizations

- Higher-order versions (Predictor–Corrector, Runga–Kutta) are called exponential integrators [Hochbruck and Lubich, 1997].
- Straightforward generalization to vector case (matrix exponential $\boldsymbol{P} = e^{t\boldsymbol{\eta}}$).
- Gaussian Quadrature with respect to weight function P.
- Conservative Exponential Integrators
- Can replace linear Green's function $e^{\eta(t-t')}$ by any $stationary$ Green's function $G(t-t')$.
- Another interesting generalization leads to Lagrangian discretizations (e.g., of the PPM type) for advection equations:

$$\frac{du}{dt} + v\frac{\partial}{\partial x}u = S(x,t,u), \qquad u(x,0) = u_0(x).$$

- η now represents the linear operator $v\frac{\partial}{\partial x}$ and $\mathcal{P}^{-1}u = e^{-tv\frac{\partial}{\partial x}}u$ corresponds to the Taylor series of $u(x-vt)$.

Charged Particle in Electromagnetic Fields

- Lorentz force:

$$\frac{m}{q}\frac{d\boldsymbol{v}}{dt} = \frac{1}{c}\boldsymbol{v}\times\boldsymbol{B} + \boldsymbol{E}.$$

- Efficiently compute the matrix exponential $\exp(\boldsymbol{\Omega})$, where

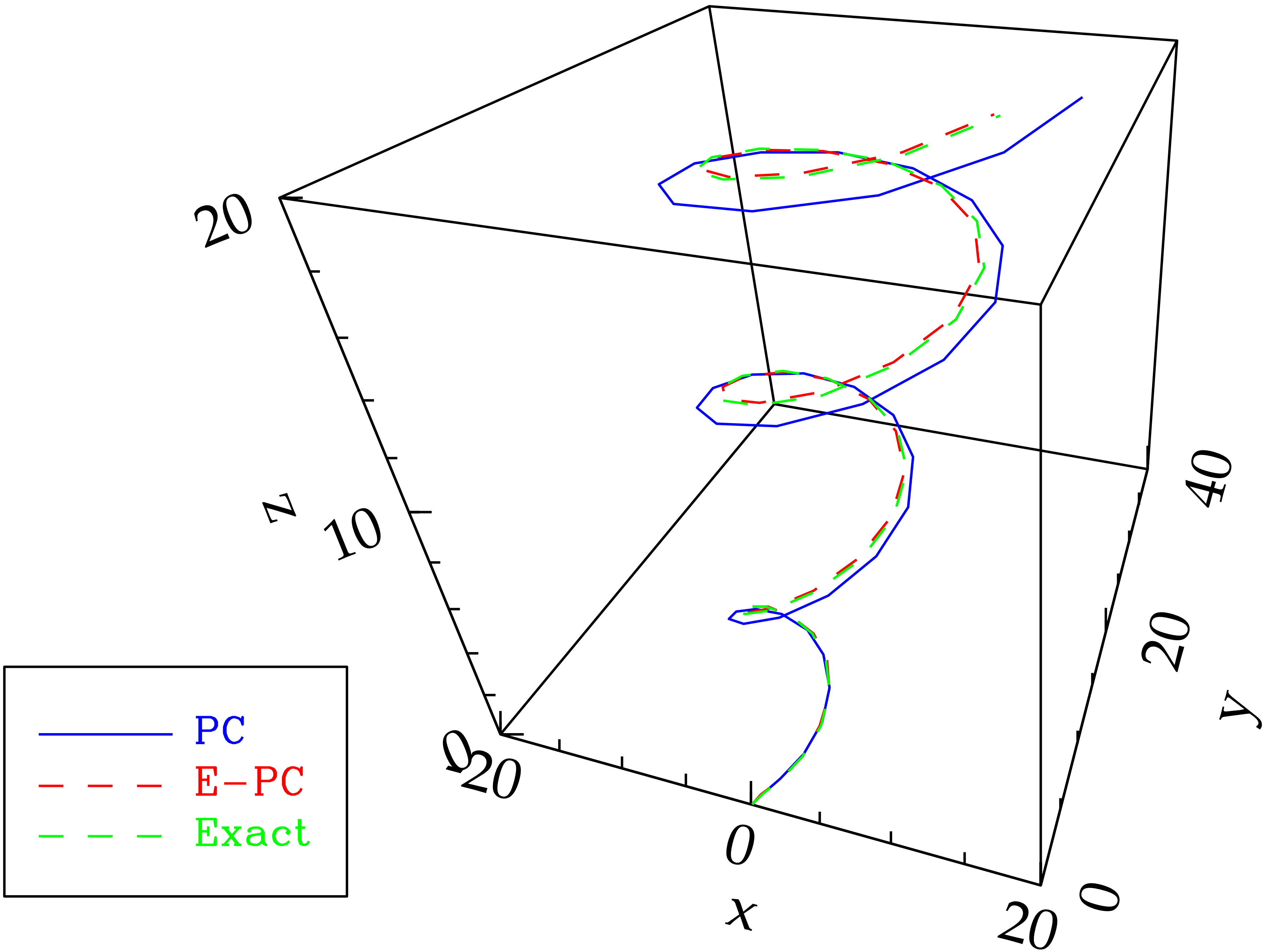
$$\boldsymbol{\Omega} = -\frac{q}{mc}t\begin{pmatrix} 0 & B_z & -B_y \\ -B_z & 0 & B_x \\ B_y & -B_x & 0 \end{pmatrix}.$$

- Requires 2 trigonometric functions, 1 division, 1 square root, and 35 additions or multiplications.

- The other necessary matrix factor, $[\exp(\boldsymbol{\Omega}) - \mathbf{1}]\boldsymbol{\Omega}^{-1}$ requires care, since $\boldsymbol{\Omega}$ is singular. Evaluate it as

$$\lim_{\lambda\to 0}[(e^{\boldsymbol{\Omega}} - \mathbf{1})(\boldsymbol{\Omega} + \lambda\mathbf{1})^{-1}].$$

Motion under Lorentz force



PC, E–PC, and exact solutions.

Conclusions

- Traditional numerical discretizations of conservative systems generically yield artificial secular drifts of nonlinear invariants.
- New exactly conservative but explicit integration algorithms have been developed.
- The transformation technique is relevant to integrable and nonintegrable Hamiltonian systems and even to non-Hamiltonian systems such as force-dissipative turbulence.
- Discretizations that preserve physically relevant structure or known analytic properties are becoming of wide interest.

References

[Bowman & Krommes 1997] J. C. Bowman & J. A. Krommes, Phys. Plasmas, **4**:3895, 1997.

[Bowman *et al.* 1993] J. C. Bowman, J. A. Krommes, & M. Ottaviani, Phys. Fluids B, **5**:3558, 1993.

[Bowman *et al.* 1997] J. C. Bowman, B. A. Shadwick, & P. J. Morrison, "Exactly conservative integrators," in *15th IMACS World Congress on Scientific Computation, Modelling and Applied Mathematics*, edited by A. Sydow, volume 2, pp. 595–600, Berlin, 1997, Wissenschaft & Technik.

[Channell & Scovel 1990] P. J. Channell & J. C. Scovel, Nonlinearity, **3**:231, 1990.

[Ge Zhong & Marsden 1988] Ge Zhong & J. E. Marsden, Phys. Lett. A, **133**:134, 1988.

[Kane *et al.* 1999] C. Kane, J. E. Marsden, & M. Ortiz, J. Math. Phys., **40**:3353, 1999.

[Kotovych & Bowman 2002] O. Kotovych & J. C. Bowman, J. Phys. A.: Math. Gen., **35**:7849, 2002.

[Ruth 1983] R. D. Ruth, IEEE Trans. Nucl. Sci., **NS-30**:2669, 1983.

[Sanz-Serna & Calvo 1994] J. M. Sanz-Serna & M. P. Calvo, *Numerical Hamiltonian Problems*, volume 7 of *Applied Mathematics and Mathematical Computation*, Chapman and Hall, London, 1994.

[Shadwick & Buell 1997] B. A. Shadwick & W. F. Buell, Phys. Rev. Lett., **79**:5189, 1997.

[Shadwick *et al.* 1999] B. A. Shadwick, J. C. Bowman, & P. J. Morrison, SIAM J. Appl.

Math., **59**:1112, 1999.