Geometric Numerical Integration of Hamiltonian systems: application to some optimal control problems

Philippe Chartier

1IPSO
INRIA-Rennes

Outline

1. First examples
   - Harmonic oscillator
   - 2-D Kepler Problem

2. Hamiltonian problems
   - Main properties of Hamiltonian systems
   - Symplectic maps
   - Application to Hamiltonian systems

3. Geometric B-series
   - B(utcher)-series
   - Algebraic characterization of geometric properties

4. Modified equations
   - Backward error analysis for ordinary differential equations
   - Geometric properties of the modified equation

5. Application to control problems
   - An optimal control problem without constraints
   - Runge-Kutta discretization of optimality conditions
Outline

1 First examples
   - Harmonic oscillator
   - 2-D Kepler Problem

2 Hamiltonian problems
   - Main properties of Hamiltonian systems
   - Symplectic maps
   - Application to Hamiltonian systems

3 Geometric B-series
   - B(utcher)-series
   - Algebraic characterization of geometric properties

4 Modified equations
   - Backward error analysis for ordinary differential equations
   - Geometric properties of the modified equation

5 Application to control problems
   - An optimal control problem without constraints
   - Runge-Kutta discretization of optimality conditions
Consider the following

**Hamiltonian**

\[ H(p, q) = \frac{1}{2}(p^2 + \omega^2 q^2) \]

and the corresponding

**Hamiltonian system**

\[
\begin{align*}
\dot{p} & = -\frac{\partial H}{\partial q} = -\omega^2 q \\
\dot{q} & = \frac{\partial H}{\partial p} = p
\end{align*}
\]

The exact trajectory is known to be an ellipse in the phase-space \((p, q)\) depending on initial conditions \((p_0, q_0)\).
Harmonic oscillator

Three elementary methods

**Explicit Euler**

\[
\begin{align*}
    p_{n+1} &= p_n + h(-\omega^2 q_n) = p_n + (-h\omega^2)q_n \\
    q_{n+1} &= q_n + h(p_n) = hp_n + q_n
\end{align*}
\]

**Implicit Euler**

\[
\begin{align*}
    p_{n+1} &= p_n + h(-\omega^2 q_{n+1}) = \frac{1}{1+h^2\omega^2} p_n + \frac{-h\omega^2}{1+h^2\omega^2} q_n \\
    q_{n+1} &= q_n + h(p_{n+1}) = \frac{h}{1+h^2\omega^2} p_n + \frac{1}{1+h^2\omega^2} q_n
\end{align*}
\]
Harmonic oscillator

Three elementary methods

Midpoint rule

\[
\begin{align*}
    p_{n+1} &= p_n + h\left(-\omega^2 \frac{q_{n+1} + q_n}{2}\right) \\
    q_{n+1} &= q_n + h\left(\frac{p_{n+1} + p_n}{2}\right)
\end{align*}
\]

i.e.

\[
\begin{align*}
    p_{n+1} &= \frac{1-h^2\omega^2/4}{1+h^2\omega^2/4}p_n + \frac{-h\omega^2}{1+h^2\omega^2/4}q_n \\
    q_{n+1} &= \frac{h}{1+h^2\omega^2/4}p_n + \frac{1-h^2\omega^2/4}{1+h^2\omega^2/4}q_n
\end{align*}
\]
Computed trajectories for the three methods

Explicit Euler (green), Midpoint Rule (red), Implicit Euler (blue)
Theoretical explanation of the different behaviors

All previous schemes can written as

A linear recurrence

\[
\begin{pmatrix}
  p_{n+1} \\
  q_{n+1}
\end{pmatrix} = M(h\omega) \begin{pmatrix}
  p_n \\
  q_n
\end{pmatrix}
\]

With, for the Explicit Euler method

\[
M(h\omega) = \begin{bmatrix}
  1 & (-h\omega^2) \\
  1 & 1
\end{bmatrix}
\]

And eigenvalues \( \lambda_{1,2} = (1 \pm ih\omega) \). Hence, the energy grows like \((1 + h^2\omega^2)^{n/2}\).
Theoretical explanation of the different behaviors

All previous schemes can be written as

a linear recurrence

\[
\begin{pmatrix}
  p_{n+1} \\
  q_{n+1}
\end{pmatrix}
= M(h\omega)
\begin{pmatrix}
  p_n \\
  q_n
\end{pmatrix}
\]

with, for the Implicit Euler method

\[
M(h\omega) = \begin{bmatrix}
  \frac{1}{1+h^2\omega^2} & \frac{-h\omega^2}{1+h^2\omega^2} \\
  \frac{h}{1+h^2\omega^2} & \frac{1}{1+h^2\omega^2}
\end{bmatrix}
\]

and eigenvalues \( \lambda_{1,2} = (1 \pm ih\omega)^{-1} \). Hence, the energy decreases like \((1 + h^2\omega^2)^{-n/2}\).
Harmonic oscillator

Theoretical explanation of the different behaviors

All previous schemes can be written as

A linear recurrence

\[
\begin{pmatrix}
  p_{n+1} \\
  q_{n+1}
\end{pmatrix}
= M(h\omega)
\begin{pmatrix}
  p_n \\
  q_n
\end{pmatrix}
\]

with, for the Midpoint rule

\[
M(h\omega) =
\begin{bmatrix}
  1 - h^2\omega^2/4 & -h\omega^2 \\
  1 + h^2\omega^2/4 & 1 + h^2\omega^2/4 \\
  h & 1 - h^2\omega^2/4 \\
  1 + h^2\omega^2/4 & 1 + h^2\omega^2/4
\end{bmatrix}
\]

and eigenvalues \(\lambda_{1,2}\) of modulus one. Hence, the energy is constant.
Consider the following Hamiltonian

\[ H(p^1, p^2, q^1, q^2) = \frac{1}{2} [(p^1)^2 + (p^2)^2] - \frac{1}{\sqrt{(q^1)^2 + (q^2)^2}}, \]

\[ = T(p) + V(q). \]

and the corresponding system

\[
\begin{align*}
\dot{p} &= -\frac{\partial H}{\partial q} = -V'(q) \\
\dot{q} &= \frac{\partial H}{\partial p} = p \quad \iff \quad \ddot{q} = -V'(q)
\end{align*}
\]

The exact trajectory is known to be an ellipse in the phase-space \((p, q)\) depending on initial conditions.
2-D Kepler Problem

Computed trajectories and energies

Euler explicit/implicit
2-D Kepler Problem

Computed trajectories and energies
2-D Kepler Problem

Computed trajectories and energies
Motivation for further investigations

Observation
Nothing as simple as a linear analysis can sustain the observed superior behavior of the midpoint rule on Kepler problem. Other non-linear problems corroborate these observations.

Consequence
A more elaborated theory is required to understand what is going on.
Outline

1. First examples
   - Harmonic oscillator
   - 2-D Kepler Problem

2. Hamiltonian problems
   - Main properties of Hamiltonian systems
   - Symplectic maps
   - Application to Hamiltonian systems

3. Geometric B-series
   - B(utcher)-series
   - Algebraic characterization of geometric properties

4. Modified equations
   - Backward error analysis for ordinary differential equations
   - Geometric properties of the modified equation

5. Application to control problems
   - An optimal control problem without constraints
   - Runge-Kutta discretization of optimality conditions
For $p$ and $q$ in $\mathbb{R}^d$, and $H$ a smooth scalar function, one can define the following Hamiltonian system:

\[
\begin{cases}
\dot{p} = -\frac{\partial H}{\partial q} \\
\dot{q} = \frac{\partial H}{\partial q}
\end{cases}
\]

Denoting

\[y = \begin{pmatrix} p \\ q \end{pmatrix}, \quad J = \begin{pmatrix} 0 & I_d \\ -I_d & 0 \end{pmatrix}\]

Canonical equations

\[\dot{y} = J^{-1}\nabla H.\]
Conservation of the Hamiltonian by the flow

**Definition**

The flow $\varphi_t$ is defined as the function which associates at time $t$ the exact solution of $\dot{y} = f(y)$ with initial condition $y(0) = y_0$.

**Theorem**

*The flow of an Hamiltonian system preserves the value of the Hamiltonian.*

**Proof**: Since $J$ is skew-symmetric, along any exact trajectory one has:

$$\frac{d}{dt} H(\varphi_t(y)) = \frac{\partial H}{\partial y} \frac{dy}{dt} = (\nabla H)^T J^{-1} \nabla H = 0.$$
Main properties of Hamiltonian systems

Conservation of volume

**Theorem**

*For a system of the form* $\dot{y} = f(y)$, *with* $\text{div} f = 0$, *one has*

$$\text{Vol}(\varphi_t(A)) = \text{Vol}(A)$$

*for any compact set* $A \subset \mathbb{R}^n$.

**Proof:** $\psi_t(y) = \frac{\partial \varphi_t}{\partial y}(y)$ *is solution of*

$$\frac{d}{dt} \psi_t(y) = \frac{\partial f}{\partial y}(\varphi_t(y))\psi_t(y), \quad \psi_0(y) = I_{\mathbb{R}^n}.$$

Hence $\frac{d}{dt} \det(\psi_t(y)) = \det(\psi_t) \text{Tr}(\psi_t^{-1}(\partial_y f(\varphi_t(y))) \psi_t) = 0$,

and $\int_{\varphi_t(A)} dz = \int_A \det(\psi_t(y)) dy = \int_A \det(\psi_0(y)) dy = \int_A dy$. 
The flow of an Hamiltonian system preserves the volume.

Proof: For an Hamiltonian system $f = J^{-1} \nabla H$

$$\text{div} f = Tr \left( \frac{\partial f}{\partial y} \right)$$

$$= Tr(J^{-1} \nabla^2 H)$$

$$= Tr(\nabla^2 H J^{-T})$$

$$= -Tr(J^{-1} \nabla^2 H)$$

$$= -\text{div} f$$

$$= 0.$$
The oriented area of the parallelogram \( P = \{ t\xi + s\eta \mid 0 \leq t, s \leq 1 \} \), generated by the two vectors 
\[
\xi = \begin{bmatrix} \xi^p \\ \xi^q \end{bmatrix} \quad \text{and} \quad \eta = \begin{bmatrix} \eta^p \\ \eta^q \end{bmatrix}
\]
is of the form:

\[
\text{oriented.area}(P) = \begin{vmatrix} \xi^p & \eta^p \\ \xi^q & \eta^q \end{vmatrix} = \xi^p \eta^q - \xi^q \eta^p.
\]
In dimension $d > 1$, one replaces this expression by the sum $\omega(\xi, \eta)$ of the oriented areas of the projections of $P$ over the $(p_i, q_i)$-planes:

$$
\omega(\xi, \eta) = \sum_{i=1}^{d} \begin{vmatrix} \xi_i^p & \eta_i^p \\ \xi_i^q & \eta_i^q \end{vmatrix} = \sum_{i=1}^{d} (\xi_i^p \eta_i^q - \xi_i^q \eta_i^p) = \xi^T J \eta.
$$
Symplectic maps

**Dimension** \( d > 1 \)

\[
\mathbb{R}^{2d-2}
\]

\[
\begin{align*}
\xi &
\end{align*}
\]

\[
\begin{align*}
\eta &
\end{align*}
\]

\[
\begin{align*}
p' &
\end{align*}
\]

\[
\begin{align*}
q' &
\end{align*}
\]

\[
(\xi_i^p \eta_i^q - \xi_i^q \eta_i^p)
\]

**Figure:** The map \( \omega \)
Symplectic maps

Linear symplectic maps

Definition

A linear map $A : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ is said to be symplectic iff:

$$A^T J A = J,$$

i.e., equivalently, iff:

$$\forall (\xi, \eta) \in \mathbb{R}^{2d} \times \mathbb{R}^{2d}, \omega(A\xi, A\eta) = \omega(\xi, \eta).$$

In dimension 1, the symplecticity of $A$ is nothing else but the preservation of areas.
Symplectic maps

Linear symplectic maps

Definition

A linear map $A : \mathbb{R}^{2d} \rightarrow \mathbb{R}^{2d}$ is said to be symplectic iff:

$$A^T J A = J,$$

i.e., equivalently, iff:

$$\forall (\xi, \eta) \in \mathbb{R}^{2d} \times \mathbb{R}^{2d}, \quad \omega(A \xi, A \eta) = \omega(\xi, \eta).$$

In dimension $d > 1$, it accounts for the preservation of the sum of the oriented areas of the projection over the $(p_i, q_i)$-planes.
Symplectic maps

General situation

Definition

A smooth map $g$ from $U$, an open subset of $\mathbb{R}^{2d}$, into $\mathbb{R}^{2d}$, is said to be symplectic if its Jacobian matrix $g'(p, q)$ is symplectic for all $(p, q)$ in $U$, i.e. iff:

$$\forall (p, q) \in U, \ (g'(p, q))^T Jg'(p, q) = J,$$

or equivalently iff:

$$\forall (p, q) \in U, \ \forall (\xi, \eta) \in \mathbb{R}^{2d} \times \mathbb{R}^{2d}, \ \omega(g'(p, q)\xi, g'(p, q)\eta) = \omega(\xi, \eta).$$
An integral quantity preserved by symplectic maps

Let $M = \psi(K)$ be a 2-D submanifold of $U$ with $\psi(s, t)$ a smooth map. $M$ can be seen as a union of small parallelograms generated by the vectors

$$\frac{\partial \psi}{\partial s} \, ds \text{ and } \frac{\partial \psi}{\partial t} \, dt,$$

and we can write the sum of oriented areas of the projections over the $(p_i, q_i)$-planes of all these parallelograms as

**The integral form of $\omega$**

$$\Omega(M) = \int \int_K \omega\left(\frac{\partial \psi}{\partial s}(s, t), \frac{\partial \psi}{\partial t}(s, t)\right) ds \, dt.$$
Theorem

Let \( g : U \rightarrow \mathbb{R}^{2d} \) be smooth and symplectic on \( U \). Then, \( g \) preserves \( \Omega(M) \), that is to say:

\[
\Omega(g(M)) = \Omega(M).
\]

Figure: Image of \( M \) by \( g \)
**Theorem**

Let $g : U \rightarrow \mathbb{R}^{2d}$ be smooth and symplectic on $U$. Then, $g$ preserves $\Omega(M)$, that is to say:

$$\Omega(g(M)) = \Omega(M).$$

Proof: $g(M)$ can be parametrized by $g \circ \psi$ on $K$ so that

$$\Omega(g(M)) = \int\int_K \omega \left( \frac{\partial g \circ \psi}{\partial s}(s, t), \frac{\partial g \circ \psi}{\partial t}(s, t) \right) ds \, dt$$

$$= \int\int_K \omega \left( g'(\psi(s, t)) \frac{\partial \psi}{\partial s}(s, t), g'(\psi(s, t)) \frac{\partial \psi}{\partial t}(s, t) \right) ds \, dt$$

$$= \int\int_K \left( \frac{\partial \psi}{\partial s}(s, t) \right)^T \left( g'(\psi(s, t)) \right)^T Jg'(\psi(s, t)) \frac{\partial \psi}{\partial t}(s, t) ds \, dt$$

$$= \Omega(M).$$
Characterization of Hamiltonian systems (I)

Theorem (Poincaré 1899)

Let $H(p, q)$ be a twice continuously differentiable function from an open subset $U \subset \mathbb{R}^{2d}$ into $\mathbb{R}^{2d}$. Then, for all $t$ such that $\varphi_t$ exists, $\varphi_t$ is symplectic.

Proof: For all $t$ such that $\varphi_t$ exists, $\Psi_t = \frac{\partial \varphi_t}{\partial y_0}$ satisfies

$$
\dot{\Psi}_t = J^{-1} \nabla^2 H(\varphi_t(y_0)) \Psi, \quad \Psi_0 = I_{2d}.
$$

Given that $\nabla^2 H(\varphi_t(y_0))$ is symmetric, we have:

$$
\frac{d}{dt} (\Psi_t^T J \Psi_t) = \left( \frac{\partial \varphi_t}{\partial y_0} \right)^T \nabla^2 H \underbrace{J^{-T} J}_{J^{-1} J} \Psi_t + \Psi_t^T J^{-1} J \nabla^2 H \Psi_t = 0.
$$

The conclusion now follows from (for $t = 0$) $\Psi_t^T J \Psi_t = I^T J I = J.$
Characterization of Hamiltonian systems (II)

**Theorem**

Let $U$ be a simply connected open subset and $f$ a smooth on $U$. If $\varphi_t(y)$ is symplectic for small $t$ and $y$ in $U$, then $\dot{y} = f(y)$ is Hamiltonian on $U$, i.e. there exists a smooth $H$ defined on $U$ such that $\forall y \in U, f(y) = J^{-1} \nabla H(y)$.

**Proof**: For all $t$ such that $\varphi_t$ exists, $\frac{\partial \varphi_t}{\partial y_0}$ is the solution of

$$
\dot{\psi}_t = f'(\varphi_t(y_0))\psi_t, \quad \psi_0 = I_{2d}.
$$

Upon differentiating the symplecticity relation, we obtain:

$$
0 = \frac{d}{dt} (\psi_t^T J \psi_t = \psi_t^T) ((f'(\varphi_t(y_0)))^T J + J^T f'(\varphi_t(y_0))) \psi_t.
$$

For $t = 0$, we get $(f'(\varphi_t(y_0)))^T J + J^T f'(\varphi_t(y_0)) = 0$, and $(Jf'(\varphi_t(y_0)))^T = Jf'(\varphi_t(y_0))$. Pfaff’s Lemma allows to conclude.
Outline

1. First examples
   - Harmonic oscillator
   - 2-D Kepler Problem

2. Hamiltonian problems
   - Main properties of Hamiltonian systems
   - Symplectic maps
   - Application to Hamiltonian systems

3. Geometric B-series
   - B(utcher)-series
   - Algebraic characterization of geometric properties

4. Modified equations
   - Backward error analysis for ordinary differential equations
   - Geometric properties of the modified equation

5. Application to control problems
   - An optimal control problem without constraints
   - Runge-Kutta discretization of optimality conditions
Given a $n$-dimensional system of differential equations

$$y'(x) = f(y(x))$$

a B-series $B(a, y)$ is a formal expression of the form

$$B(a) = \text{id}_{\mathbb{R}^n} + \sum_{t \in \mathcal{T}} \frac{h^{\lvert t \rvert}}{\sigma(t)} a(t) F(t)$$

$$= \text{id}_{\mathbb{R}^n} + h a(\cdot) f(\cdot) + h^2 a(\mathcal{J})(f'f)(\cdot) + \cdots$$

- the index set $\mathcal{T} = \{\cdot, \mathcal{J}, \mathcal{V}, \mathcal{L}, \cdots\}$ is the set of trees,
- $\lvert t \rvert$ and $\sigma(t)$ are fixed positive integers,
- $F(t)$ is a map from $\mathbb{R}^n$ to $\mathbb{R}^n$ obtained from $f$ and its derivatives,
- $a$ is a function defined on $\mathcal{T}$ which characterizes $B(a, y)$. 
Taylor series of the Euler approximation

In trying to get the Taylor expansion of the Implicit Euler approximation

\[ y_1 = y_0 + hf(y_1) \]

one gets successively

\[
\begin{align*}
y_1 & = y_0 + h f + \mathcal{O}(h^2), \\
& = y' \\
y_1 & = y_0 + h f + h^2 f' f + \mathcal{O}(h^3), \\
& = y' \\
y_1 & = y_0 + h f + h^2 f' f + h^3 \left( f' f + \frac{1}{2} f'' (f, f) \right) + \mathcal{O}(h^4). \\
& \neq y^{(3)} = f' f + f'' (f, f)
\end{align*}
\]
Rooted trees

Definition (Rooted trees)

The set of rooted trees is recursively defined by:

1. \( \bullet \in T \), \( \sigma(\bullet) = 1 \).
2. if \((t_1, \ldots, t_n) \in T^n \) are distinct trees then \( t = [t_1, \ldots, t_1, \ldots, t_n, \ldots, t_n] \in T \) and \( \sigma(t) = \prod_{i=1}^{n} r_i! \sigma(t_i)^{r_i} \).

The order of a tree \(|t|\) is its number of vertices.

Example

<table>
<thead>
<tr>
<th>Tree ( t )</th>
<th>( \bullet )</th>
<th>( \bullet )</th>
<th>( \bullet )</th>
<th>( \bullet )</th>
<th>( \bullet )</th>
<th>( \bullet )</th>
<th>( \bullet )</th>
<th>( \bullet )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order (</td>
<td>t</td>
<td>)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Symmetry ( \sigma(t) )</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>
Elementary differentials

Definition
For each \( t \in \mathcal{T} \), the elementary differential \( F(t) \) associated with \( t \) is the mapping from \( \mathbb{R}^n \) to \( \mathbb{R}^n \), defined recursively by:

1. \( F(\cdot)(y) = f(y) \),
2. \( F([t_1, \ldots, t_n])(y) = f^{(n)}(y) \left( F(t_1)(y), \ldots, F(t_n)(y) \right) \).

Example

\[
\begin{align*}
F(\cdot)(y) &= f'(y)f(y), \\
F(\cdot{\cdot})(y) &= f'(y)f'(y)f(y), \\
F(\cdot{\cdot}{\cdot})(y) &= f^{(3)}(y) \left( f(y), f(y), f(y) \right).
\end{align*}
\]
B-series integrators

B-series expansions of some integrators

1. Exact solution: \( y(h) = y + hf() + \frac{h^2}{2.1} (f'f)(y) + \frac{h^3}{3.2} (f''(f, f))(y) + \frac{h^3}{6.1} (f'f'f)(y) + \ldots = B(1/\gamma, y_0) \) with \( \gamma([t_1, \ldots, t_n] = |t|\gamma(t_1) \cdots \gamma(t_n). \)
B-series integrators

B-series expansions of some integrators

1. **Exact solution:**
   \[ y(h) = y + hf() + \frac{h^2}{2.1} (f'f)(y) + \frac{h^3}{3.2} (f''(f, f))(y) + \frac{h^3}{6.1} (f'f'f)(y) + \ldots = B(1/\gamma, y_0) \text{ with } \gamma([t_1, \ldots, t_n] = |t| \gamma(t_1) \cdots \gamma(t_n).) \]

2. **Explicit Euler:**
   \[ y + hf(y) = B(a, y) \text{ with } a(\cdot) = 1 \text{ and } a(t) = 0 \text{ for all } t \neq \cdot. \]
B-series integrators

B-series expansions of some integrators

1. Exact solution: $y(h) = y + hf() + \frac{h^2}{2.1} (f' f)(y) + \frac{h^3}{3.2} (f'' (f, f))(y) + \frac{h^3}{6.1} (f' f' f)(y) + \ldots = B(1/\gamma, y_0)$ with $\gamma([t_1, \ldots, t_n] = \vert t \vert \gamma(t_1) \cdots \gamma(t_n)$.

2. Explicit Euler: $y + hf(y) = B(a, y)$ with $a(\cdot) = 1$ and $a(t) = 0$ for all $t \neq \cdot$.

3. Implicit Euler: $Y = y + hf(Y)$ and $y + hf(Y) = B(a, y)$ with $a(t) = 1$ for all $t \in T$. 
B-series expansions of some integrators

1. **Exact solution**: \( y(h) = y + hf() + \frac{h^2}{2.1} (f'f)(y) + \frac{h^3}{3.2} (f''(f, f))(y) + \frac{h^3}{6.1} (f'f'f))(y) + \ldots = B(1/\gamma, y_0) \) with \( \gamma([t_1, \ldots, t_n] = |t|\gamma(t_1) \cdots \gamma(t_n). \)

2. **Explicit Euler**: \( y + hf(y) = B(a, y) \) with \( a(\cdot) = 1 \) and \( a(t) = 0 \) for all \( t \neq \cdot \).

3. **Implicit Euler**: \( Y = y + hf(Y) \) and \( y + hf(Y) = B(a, y) \) with \( a(t) = 1 \) for all \( t \in T \).

4. **Midpoint rule**: \( Y = y + \frac{h}{2} f(Y) \) and \( y + hf(Y) = B(a, y) \) with \( a(t) = (1/2)|t|^{-1} \) for all \( t \in T \).
Assume \( l(\varphi_t(y)) = l(y) \) for all \((t, y)\). The integrator \( B(a, y) \) preserves \( l \) iff for all \( y \), \( l\left(B(a, y)\right) = l(y) \).

**Theorem**

The integrator \( B(a, y) \) preserves \( l \) that for all couples \((f, l)\) of a vector field \( f \) and a first integral \( l \), iff \( a(\cdot) = 1 \) and \( a \) satisfies

\[
\forall m \geq 2, \forall (t_1, \ldots, t_m) \in T^m, \quad a(t_1) \cdot \cdots \cdot a(t_m) = \sum_{j=1}^{m} a(t_j \circ \prod_{i \neq j} t_i).
\]

The notation \( s \circ \prod_{j} t_j \) is used here to denote the tree obtained by connecting the roots of all trees \( t_j \) to the root of \( s \).
Theorem

Consider a B-series integrator $B(a, y)$. Then,

- it can not preserve general invariants for all problems $(f, I)$.
- it preserves quadratic invariants for all problems $(f, I)$ with $I(y) = y^T Cy$ iff

  $$\forall (t_1, t_2) \in T^2, \quad a(t_1)a(t_2) = a(t_1 \circ t_2) + a(t_2 \circ t_1).$$  \hfill (1)

- it is symplectic for all problems $f = J^{-1}\nabla H$ iff (1) holds true.
- it preserves $H$ for all problems $f = J^{-1}\nabla H$ only if it is conjugate to a symplectic B-series integrator.
- if it is symplectic, it preserves a modified Hamiltonian for all problems $f = J^{-1}\nabla H$.
- it can not be symplectic and preserve $H$ for all problems.
Outline

1. First examples
   - Harmonic oscillator
   - 2-D Kepler Problem

2. Hamiltonian problems
   - Main properties of Hamiltonian systems
   - Symplectic maps
   - Application to Hamiltonian systems

3. Geometric B-series
   - B(utcher)-series
   - Algebraic characterization of geometric properties

4. Modified equations
   - Backward error analysis for ordinary differential equations
   - Geometric properties of the modified equation

5. Application to control problems
   - An optimal control problem without constraints
   - Runge-Kutta discretization of optimality conditions
Concept of backward error analysis

Given a differential equation

\[ \dot{y} = f(y), \quad y(0) = y_0 \]

and a numerical one-step method (say a B-series integrator)

\[ y_{n+1} = B(a, y) \]

the idea of **backward analysis** is to find a modified differential equation

\[ \hat{y} = \hat{f}_h(\tilde{y}) = f(\tilde{y}) + hf_2(\tilde{y}) + h^2f_3(\tilde{y}) + h^3f_4(\tilde{y}) + \ldots, \quad \tilde{y}(0) = y_0 \]

such that for \( t_n = nh \)

\[ y_n = \tilde{y}(t_n) \]
Backward error analysis for ordinary differential equations

\[ \dot{y} = f(y) \]

\[ \dot{z} = f_h(z) \]

\[ y_0, y_1, y_2, y_3, \ldots = \]

\[ z(0), z(h), z(2h), \ldots \]
Backward error analysis for ordinary differential equations

Illustration with a toy-problem

**Example**

Lotka-Volterra in normal form

\[
\begin{align*}
\dot{u} &= e^v - 2 \\
\dot{v} &= 1 - e^u
\end{align*}
\]

i.e. \( y' = f(y) \) with \( f(y) = (e^v - 2, 1 - e^u)^T \).

**Elementary differentials**

\[
F(\dot{\cdot}) = f'f = \begin{pmatrix} e^v(1 - e^u) \\ -e^u(e^v - 2) \end{pmatrix},
F(\ddot{\cdot}) = f'f'f = -e^{u+v}f,
F(\dddot{\cdot}) = f'''(f, f) = \begin{pmatrix} e^v(1 - e^u)^2 \\ -e^u(e^v - 2)^2 \end{pmatrix}.
\]
Backward error analysis for ordinary differential equations

**Modified vector field**

\[
\tilde{f}_h = f + h \left( a(\dot{x}) - \frac{1}{2} \right) f' f + \frac{h^2}{2} \left( a(\ddot{x}) - a(\dot{x}) + \frac{1}{6} \right) f''(f, f) \\
+ h^2 \left( a(\ddot{x}) - a(\dot{x}) + \frac{1}{3} \right) f' f' f + \mathcal{O}(h^3).
\]

**Explicit Euler method**

\[
\tilde{f}_{he} = f - \frac{1}{2} h f' f + \frac{h^2}{12} f''(f, f) + \frac{h^2}{3} f' f' f + \mathcal{O}(h^3),
\]
Modified vector field

\[ \tilde{f}_h = f + h \left( a(\dot{x}) - \frac{1}{2} \right) f' f + \frac{h^2}{2} \left( a(\ddot{x}) - a(\dot{x}) + \frac{1}{6} \right) f''(f, f) + h^2 \left( a(\dddot{x}) - a(\ddot{x}) + \frac{1}{3} \right) f' f' f + O(h^3). \]

Implicit Euler method

\[ \tilde{f}_{iE}^h = f + \frac{1}{2} h f' f + \frac{h^2}{12} f''(f, f) + \frac{h^2}{3} f' f' f + O(h^3), \]
Modified vector field

\[ \tilde{f}_h = f + h \left( a(\dot{x}) - \frac{1}{2} \right) f'f + \frac{h^2}{2} \left( a(\ddot{x}) - a(\dot{x}) + \frac{1}{6} \right) f''(f, f) + h^2 \left( a(\dddot{x}) - a(\ddot{x}) + \frac{1}{3} \right) f'f'f + O(h^3). \]

Midpoint Rule

\[ \tilde{f}^m_{th} = f - \frac{h^2}{24} f''(f, f) + \frac{h^2}{12} f'f'f + O(h^3). \]
Figure: Exact solutions of modified equations (red) versus Explicit Euler (green).
Figure: Exact solutions of modified equations (red) versus Implicit Euler (green).
Figure: Exact solutions of modified equations (red) versus Midpoint Rule (green).
Theorem

The modified vector field $\tilde{f}_h$ of an integrator $B(a, y)$ satisfies:

- **If $\Phi_{f,h}$ has order $p$, then**

  $$\tilde{f}_h(y) = f(y) + h^p f_{p+1}(y) + h^{p+1} f_{p+2}(y) + \ldots.$$ 

- **If $B(a, y)$ is symmetric, then $f_{2j} = 0$ for all $j$.**

- **If $B(a, y)$ preserves a first integral $l(y)$, then $l(y)$ is a first integral of the modified differential equation.**

- **If $B(a, y)$ is symplectic for $f(y) = J^{-1} \nabla H(y)$, then $\tilde{f}_h$ is Hamiltonian.**

- **If $B(a, y)$ is volume-preserving for a divergence free $f$, then $\tilde{f}_h$ is divergence free.**
Outline

1. First examples
   - Harmonic oscillator
   - 2-D Kepler Problem

2. Hamiltonian problems
   - Main properties of Hamiltonian systems
   - Symplectic maps
   - Application to Hamiltonian systems

3. Geometric B-series
   - B(utcher)-series
   - Algebraic characterization of geometric properties

4. Modified equations
   - Backward error analysis for ordinary differential equations
   - Geometric properties of the modified equation

5. Application to control problems
   - An optimal control problem without constraints
   - Runge-Kutta discretization of optimality conditions
An optimal control problem without constraints

Original formulation

\[(P)\] \[
\begin{align*}
\text{Find } (y, u) \text{ satisfying} \\
&\min \Phi(y(1)), \\
&\dot{y}(t) = f(y(t), u(t)), \quad t \in (0, 1), \\
&y(0) = y^0.
\end{align*}
\]

Pontryagin formulation \((H(y, p, u) := p^T f(y, u))\)

\[(OC)\] \[
\begin{align*}
\min \Phi(y(1)), \\
\dot{y}(t) &= H_p(y(t), p(t), u(t)) \\
\dot{p}(t) &= -H_y(y(t), p(t), u(t)) \\
H(y(t), p(t), u(t)) &= \min_\alpha H(y(t), p(t), \alpha) \\
y(0) &= y^0, \quad p(1) = \Phi'(y(1)).
\end{align*}
\]
An optimal control problem without constraints

**Original formulation**

\[
(P) \begin{cases}
  \text{Find } (y, u) \text{ satisfying } \\
  \min \Phi(y(1)), \\
  \dot{y}(t) = f(y(t), u(t)), \quad t \in (0, 1), \\
  y(0) = y^0. 
\end{cases}
\]

**Hamiltonian formulation** \( (OC') \)

\[
(OC') \begin{cases}
  \min \Phi(y(1)), \\
  \dot{y}(t) = \mathcal{H}_p(y(t), p(t)) \\
  \dot{p}(t) = -\mathcal{H}_y(y(t), p(t)) \\
  y(0) = y^0, \quad p(1) = \Phi'(y(1)).
\end{cases}
\]
The Runge-Kutta discretisation $(A, b)$ for problem $(P)$

\[
(DP) \begin{cases} 
\min \Phi(y_N), \\
y_{k+1} = y_k + h \sum_{i=1}^{s} b_i f(y_{ki}, u_{ki}), & k = 0, \ldots, N - 1, \\
y_{ki} = y_k + h \sum_{j=1}^{s} a_{ij} f(y_{kj}, u_{kj}), & i = 1, \ldots, s, \\
y_0 = y^0.
\end{cases}
\]

is equivalent Hager[00] or Bonnans/Laurent-Varin[06] to

the symplectic partitionned RK-discretisation for $(OC')$

\[
(DOC) \begin{cases} 
\min \Phi(y_N), \\
y_{k+1} = y_k + h \sum_{i=1}^{s} b_i \mathcal{H}_p(y_{ki}, p_{ki}), \\
y_{ki} = y_k + h \sum_{j=1}^{s} a_{ij} \mathcal{H}_p(y_{kj}, p_{kj}), & i = 1, \ldots, s, \\
p_{k+1} = p_k - h \sum_{i=1}^{s} b_i \mathcal{H}_y(y_{ki}, p_{ki}), \\
p_{ki} = p_k - h \sum_{j=1}^{s} \hat{a}_{ij} \mathcal{H}_y(y_{kj}, p_{kj}), & i = 1, \ldots, s, \\
y_0 = y^0, & p_N = \Phi'(y_N).
\end{cases}
\]
Modified optimal control problem

Existence of a modified problem

**Question**

Can one interpret the numerical solution obtained as the exact solution of a modified optimal control problem of the form:

\[
(P) \left\{ \begin{array}{c}
\text{Min } \Phi(x(1)), \\
\dot{x}(t) = \tilde{f}(x(t), u(t)), \quad t \in (0, 1), \\
x(0) = x^0,
\end{array} \right.
\]

where

\[
\tilde{f}(x, u) = f(x, u) + hf_2(x, u) + h^2f_3(x, u) + \ldots?
\]
Modified optimal control problem

Note that if this is possible, then problem \((\tilde{P})\) has necessary conditions \((\tilde{OC})\) from the Pontryagin principle,

\[
\begin{aligned}
\dot{x}(t) &= \tilde{f}(x(t), u(t)), \\
\dot{p}(t) &= -p^T\tilde{f}_x(x(t), u(t)), \\
0 &= p^T\tilde{f}_u(x(t), u(t)) \iff u(t) = \tilde{\varphi}(x(t), p(t)), \\
x(0) = x^0, & p(1) = \Phi'(x(1)),
\end{aligned}
\]

where

\[
\tilde{\varphi}(x, p) = \varphi(x, p) + h\varphi_2(x, p) + h^2\varphi_3(x, p) + \ldots.
\]

is (formally) given by

\[
0 = p^T\tilde{f}_u(x, u) \iff u = \tilde{\varphi}(x, p).
\]
Modified optimal control problem

The linear case

Equations \((x, u \in \mathbb{R}^n, A, Z, S, B \in \mathbb{R}^{n \times n}, Z^T = Z, S^T = S, \det(B) \neq 0)\)

\[
\begin{align*}
\min & \quad \frac{1}{2} \int_0^1 (x^T Z x + u^T S u) dt, \\
\dot{x} & = Ax + Bu \\
x(0) & \text{ given}
\end{align*}
\]

or equivalently:

Hamiltonian system \((H = \frac{1}{2} (x^T Z x + u^T S u) + p^T (Ax + Bu))\)

\[
\dot{y} = \begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} A & -BS^{-1}B^T \\ -Z & -A^T \end{pmatrix} y \ := \ My.
\]
Symplectic Runge-Kutta discretization

\[ \dot{y} = M_h y, \]

where

\[ M_h = \begin{pmatrix} A_h & -\Lambda_h \\ -Z_h & -A_h^T \end{pmatrix}, \]

and \( \Lambda_h = BS^{-1}B^T + O(h) > 0 \) and symmetric for small \( h \). Thus:

\[ \Lambda_h = B_h S^{-1} B_h^T. \]

Modified optimal control problem

\[
\begin{aligned}
\begin{cases}
\text{Min} & \frac{1}{2} \int_0^1 (x^T Z_h z + u^T S u) dt \\
\dot{x} & = A_h x + B_h u \\
x(0) & \text{given}
\end{cases}
\end{aligned}
\]

When \( \text{dim} u < \text{dim} x \), one can add extra control variables and recover the result (though for a stationary problem).
These preliminary results raise several questions:

1. Modified control problem: is it possible to extend to non-linear problems, using the formalism of B-series?

2. What is the advantage of symplecticity for optimal control problems? Can this be seen from the modified problem, as this is the case for ordinary differential equations?

3. Conjugate points can be defined for stationary point control problems. Computing numerically the conjugate points with the criteria $\det \frac{\partial x_n}{\partial p_0} = 0$ (using the variational equation) yields exactly the conjugate points of the modified control problem... Can symplectic methods be useful for computing conjugate points?