Numerical Solution of Differential Equations I

Endre Süli

Mathematical Institute
University of Oxford
2020

Lecture 11

Motivation

Example (MATLAB)

Compare the implicit midpoint rule with the explicit and implicit Euler methods for the following initial-value problem:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \qquad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Motivation

Example (MATLAB)

Compare the implicit midpoint rule with the explicit and implicit Euler methods for the following initial-value problem:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \qquad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Exact solution: $y_1(t) = \sin t$, $y_2(t) = \cos t$.

Motivation

Example (MATLAB)

Compare the implicit midpoint rule with the explicit and implicit Euler methods for the following initial-value problem:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \qquad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Exact solution: $y_1(t) = \sin t$, $y_2(t) = \cos t$.

Clearly,

$$Q(t) := \sqrt{y_1^2(t) + y_2^2(t)} = 1 \qquad \text{for all } t \geq 0.$$

Motivation

Example (MATLAB)

Compare the implicit midpoint rule with the explicit and implicit Euler methods for the following initial-value problem:

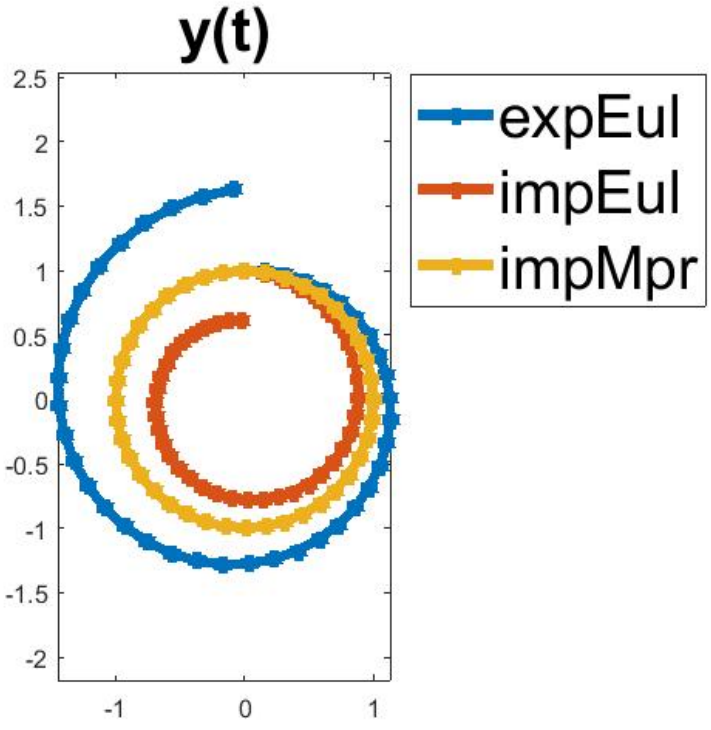
$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \qquad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

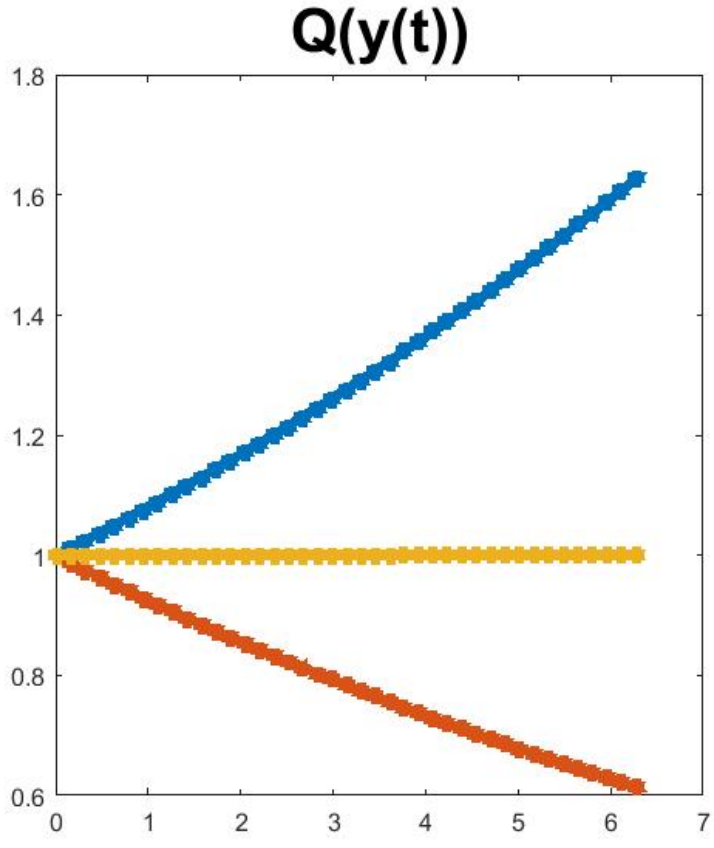
Exact solution: $y_1(t) = \sin t$, $y_2(t) = \cos t$.

Clearly,

$$Q(t) := \sqrt{y_1^2(t) + y_2^2(t)} = 1 \qquad \text{for all } t \geq 0.$$

[Run the MATLAB code: testcase2a.m]


y(t)
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-1
0
1
expEul
impEul
impMpr


Q(y(t))
1.8
1.6
1.4
1.2
1
0.8
0.6
0
1
2
3
4
5
6
7

The implicit midpoint rule for $\mathbf{y}'(t) = \mathbf{f}(\mathbf{y}(t))$, $\mathbf{y}(0) = \mathbf{x}$

$$\frac{\mathbf{y}_{n+1} - \mathbf{y}_n}{h} = \mathbf{f}\left(\frac{\mathbf{y}_{n+1} + \mathbf{y}_n}{2}\right), \quad n = 0, 1, \ldots,$$

$$\mathbf{y}_0 = \mathbf{x} = \text{given}.$$

The implicit midpoint rule for $\mathbf{y}'(t) = \mathbf{f}(\mathbf{y}(t))$, $\mathbf{y}(0) = \mathbf{x}$

$$\frac{\mathbf{y}_{n+1} - \mathbf{y}_n}{h} = \mathbf{f}\left(\frac{\mathbf{y}_{n+1} + \mathbf{y}_n}{2}\right), \quad n = 0, 1, \ldots,$$

$$\mathbf{y}_0 = \mathbf{x} = \text{given}.$$

Exercise. Let $(\cdot, \cdot)$ be the inner product in $\mathbb{R}^d$ and let $\|\cdot\|$ be the associated Euclidean norm in $\mathbb{R}^d$. Suppose that $\mathbf{f} : D \subset \mathbb{R}^d \to \mathbb{R}^d$ and $(\mathbf{f}(\mathbf{z}), \mathbf{z}) = 0$ for all $\mathbf{z} \in D$. Show that if $\mathbf{y}' = \mathbf{f}(\mathbf{y})$, $\mathbf{y}(0) = \mathbf{x}$ and $\mathbf{y}_n$ is the implicit midpoint approximation of $\mathbf{y}$, then:

(a) $\|\mathbf{y}(t)\|^2 = \|\mathbf{x}\|^2$ for all $t \geq 0$;

(b) $\|\mathbf{y}_n\|^2 = \|\mathbf{x}\|^2$ for all $n \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

The implicit midpoint rule for $\mathbf{y}'(t) = \mathbf{f}(\mathbf{y}(t))$, $\mathbf{y}(0) = \mathbf{x}$

$$\frac{\mathbf{y}_{n+1} - \mathbf{y}_n}{h} = \mathbf{f}\left(\frac{\mathbf{y}_{n+1} + \mathbf{y}_n}{2}\right), \quad n = 0, 1, \ldots,$$

$$\mathbf{y}_0 = \mathbf{x} = \text{given}.$$

Exercise. Let $(\cdot, \cdot)$ be the inner product in $\mathbb{R}^d$ and let $\|\cdot\|$ be the associated Euclidean norm in $\mathbb{R}^d$. Suppose that $\mathbf{f} : D \subset \mathbb{R}^d \to \mathbb{R}^d$ and $(\mathbf{f}(\mathbf{z}), \mathbf{z}) = 0$ for all $\mathbf{z} \in D$. Show that if $\mathbf{y}' = \mathbf{f}(\mathbf{y})$, $\mathbf{y}(0) = \mathbf{x}$ and $\mathbf{y}_n$ is the implicit midpoint approximation of $\mathbf{y}$, then:

(a) $\|\mathbf{y}(t)\|^2 = \|\mathbf{x}\|^2$ for all $t \geq 0$;

(b) $\|\mathbf{y}_n\|^2 = \|\mathbf{x}\|^2$ for all $n \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

Hint: Take the inner product of
(a) the ODE with $\mathbf{y}(t)$; and
(b) the implicit midpoint rule with $\frac{\mathbf{y}_{n+1}+\mathbf{y}_n}{2}$.

Structure-preserving integrators

In many time-dependent physical problems modelled by initial-value problems, certain relevant physical quantities (e.g. energy, mass, volume, etc.), are preserved during the course of evolution, i.e. they are constant in time.

Structure-preserving integrators

In many time-dependent physical problems modelled by initial-value problems, certain relevant physical quantities (e.g. energy, mass, volume, etc.), are preserved during the course of evolution, i.e. they are constant in time.

One needs to preserve some of these quantities at the discrete level.

Structure-preserving integrators

In many time-dependent physical problems modelled by initial-value problems, certain relevant physical quantities (e.g. energy, mass, volume, etc.), are preserved during the course of evolution, i.e. they are constant in time.

One needs to preserve some of these quantities at the discrete level.

For simplicity, we restrict ourselves to the autonomous ODE

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}), \quad \text{where} \quad \mathbf{f} : D \to \mathbb{R}^d, \tag{1}$$

(where now $\mathbf{y}$ is considered to be a function of $t \in [0, \infty)$, and $\mathbf{y}' := \mathrm{d}\mathbf{y}/\mathrm{d}t$), subject to the initial condition

$$\mathbf{y}(0) = \mathbf{x},$$

where $\mathbf{x} \in D$, and D is a nonempty open subset of $\mathbb{R}^d$.

Definition

For $t \geq 0$, let $\mathbf{\Phi}^t : D \to \mathbb{R}^d$ denote the function that maps an initial datum $\mathbf{x} \in D$ into $\mathbf{y}(t) \in \mathbb{R}^d$, where $\mathbf{y}(t)$ is the solution at time t to $\mathbf{y}' = \mathbf{f}(\mathbf{y})$, $\mathbf{y}(0) = \mathbf{x}$ (tacitly assuming that the solution $t \in [0, \infty) \mapsto \mathbf{y}(t) \in \mathbb{R}^d$ to this initial-value problem, for each $\mathbf{x} \in D$, exists and that it is unique).

Definition

For $t \geq 0$, let $\mathbf{\Phi}^t : D \to \mathbb{R}^d$ denote the function that maps an initial datum $\mathbf{x} \in D$ into $\mathbf{y}(t) \in \mathbb{R}^d$, where $\mathbf{y}(t)$ is the solution at time t to $\mathbf{y}' = \mathbf{f}(\mathbf{y})$, $\mathbf{y}(0) = \mathbf{x}$ (tacitly assuming that the solution $t \in [0, \infty) \mapsto \mathbf{y}(t) \in \mathbb{R}^d$ to this initial-value problem, for each $\mathbf{x} \in D$, exists and that it is unique).

The family $\{\mathbf{\Phi}^t\}_{t \geq 0}$ is called the *flow* of (1) (defined on $D \subset \mathbb{R}^d$).

Remark

The function $t \mapsto \mathbf{\Phi}^t \mathbf{x}$ *is the solution to* $\mathbf{y}' = \mathbf{f}(\mathbf{y})$, $\mathbf{y}(0) = \mathbf{x}$.

Using the concept of flow, we can clarify what is a “preserved quantity”.

Using the concept of flow, we can clarify what is a "preserved quantity".

Definition

A *first integral* of (1) is a function $I : D \to \mathbb{R}$ that satisfies $I(\mathbf{\Phi}^t \mathbf{x}) = I(\mathbf{x})$ for every $\mathbf{x} \in D$ and every admissible $t \geq 0$.

Lemma

I is a first integral of (1) if, and only if, $\frac{\mathrm{d}}{\mathrm{d}t} I(\mathbf{\Phi}^t \mathbf{x}) = 0$ for every $\mathbf{x} \in D$ and every admissible $t \geq 0$. This is equivalent to:

$$\mathbf{D}I(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x}) = 0 \qquad \text{for every } \mathbf{x} \in D,$$

where $\mathbf{D}I(\mathbf{x}) := \mathbf{grad}\, I(\mathbf{x}) = (\frac{\partial I}{\partial x_1}, \ldots, \frac{\partial I}{\partial x_d})^{\mathrm{T}}$.

Lemma

I is a first integral of (1) if, and only if, $\frac{\mathrm{d}}{\mathrm{d}t} I(\mathbf{\Phi}^t \mathbf{x}) = 0$ for every $\mathbf{x} \in D$ and every admissible $t \geq 0$. This is equivalent to:

$$\mathbf{D}I(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x}) = 0 \qquad \text{for every } \mathbf{x} \in D,$$

where $\mathbf{D}I(\mathbf{x}) := \mathbf{grad}\, I(\mathbf{x}) = (\frac{\partial I}{\partial x_1}, \ldots, \frac{\partial I}{\partial x_d})^{\mathrm{T}}$.

Proof. The first part of the lemma is trivial.

Lemma

I is a first integral of (1) if, and only if, $\frac{\mathrm{d}}{\mathrm{d}t} I(\mathbf{\Phi}^t \mathbf{x}) = 0$ for every $\mathbf{x} \in D$ and every admissible $t \geq 0$. This is equivalent to:

$$\mathbf{D}I(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x}) = 0 \qquad \text{for every } \mathbf{x} \in D,$$

where $\mathbf{D}I(\mathbf{x}) := \mathbf{grad}\, I(\mathbf{x}) = (\frac{\partial I}{\partial x_1}, \ldots, \frac{\partial I}{\partial x_d})^{\mathrm{T}}$.

Proof. The first part of the lemma is trivial. The second part follows by applying the chain rule on the left-hand side of the equality $\frac{\mathrm{d}}{\mathrm{d}t} I(\mathbf{\Phi}^t(\mathbf{x})) = 0$.

Lemma

I is a first integral of (1) if, and only if, $\frac{\mathrm{d}}{\mathrm{d}t} I(\mathbf{\Phi}^t \mathbf{x}) = 0$ for every $\mathbf{x} \in D$ and every admissible $t \geq 0$. This is equivalent to:

$$\mathbf{D}I(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x}) = 0 \qquad \text{for every } \mathbf{x} \in D,$$

where $\mathbf{D}I(\mathbf{x}) := \mathbf{grad}\, I(\mathbf{x}) = (\frac{\partial I}{\partial x_1}, \ldots, \frac{\partial I}{\partial x_d})^{\mathrm{T}}$.

Proof. The first part of the lemma is trivial. The second part follows by applying the chain rule on the left-hand side of the equality $\frac{\mathrm{d}}{\mathrm{d}t} I(\mathbf{\Phi}^t(\mathbf{x})) = 0$. Indeed, for all $t \geq 0$,

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} I(\mathbf{\Phi}^t(\mathbf{x})) = \frac{\mathrm{d}}{\mathrm{d}t} I(\mathbf{y}(t)) = \mathbf{D}I(\mathbf{y}(t)) \cdot \mathbf{y}'(t) = \mathbf{D}I(\mathbf{y}(t)) \cdot \mathbf{f}(\mathbf{y}(t)),$$

where $\mathbf{y}$ is the solution of the initial-value problem $\mathbf{y}'(t) = \mathbf{f}(\mathbf{y}(t))$, $\mathbf{y}(0) = \mathbf{x} \in D$.

Lemma

I is a first integral of (1) if, and only if, $\frac{\mathrm{d}}{\mathrm{d}t} I(\mathbf{\Phi}^t \mathbf{x}) = 0$ for every $\mathbf{x} \in D$ and every admissible $t \geq 0$. This is equivalent to:

$$\mathbf{D}I(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x}) = 0 \qquad \text{for every } \mathbf{x} \in D,$$

where $\mathbf{D}I(\mathbf{x}) := \mathbf{grad}\, I(\mathbf{x}) = (\frac{\partial I}{\partial x_1}, \ldots, \frac{\partial I}{\partial x_d})^{\mathrm{T}}$.

Proof. The first part of the lemma is trivial. The second part follows by applying the chain rule on the left-hand side of the equality $\frac{\mathrm{d}}{\mathrm{d}t} I(\mathbf{\Phi}^t(\mathbf{x})) = 0$. Indeed, for all $t \geq 0$,

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} I(\mathbf{\Phi}^t(\mathbf{x})) = \frac{\mathrm{d}}{\mathrm{d}t} I(\mathbf{y}(t)) = \mathbf{D}I(\mathbf{y}(t)) \cdot \mathbf{y}'(t) = \mathbf{D}I(\mathbf{y}(t)) \cdot \mathbf{f}(\mathbf{y}(t)),$$

where $\mathbf{y}$ is the solution of the initial-value problem $\mathbf{y}'(t) = \mathbf{f}(\mathbf{y}(t))$, $\mathbf{y}(0) = \mathbf{x} \in D$. Thus in particular $0 = \mathbf{D}I(\mathbf{y}(0)) \cdot \mathbf{f}(\mathbf{y}(0))$, and the assertion follows, because $\mathbf{y}(0) = \mathbf{x} \in D$. $\diamond$

For a systematic investigation, we consider first integrals that can be expressed as polynomials.

For a systematic investigation, we consider first integrals that can be expressed as polynomials.

Definition

We shall say that a first integral I of an autonomous system is a polynomial of degree $n \in \mathbb{N}$ if

$$I(\mathbf{x}) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}_0^d,\, |\boldsymbol{\alpha}| \leq n} \beta_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}, \tag{2}$$

where $\beta_{\boldsymbol{\alpha}} \in \mathbb{R}$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$, $|\boldsymbol{\alpha}| = \sum_{i=1}^d \alpha_i$, and $\mathbf{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$; in other words, I is a multivariate polynomial of degree n in $\mathbf{x} \in \mathbb{R}^d$.

Example

Linear first integrals are of the form $I(\mathbf{x}) = \mathbf{b}^{\mathrm{T}}\mathbf{x} + c$ (with $\mathbf{b} \in \mathbb{R}^d$, and $c \in \mathbb{R}$).

Example

Linear first integrals are of the form $I(\mathbf{x}) = \mathbf{b}^{\mathrm{T}}\mathbf{x} + c$ (with $\mathbf{b} \in \mathbb{R}^d$, and $c \in \mathbb{R}$).

Example

Quadratic first integrals are of the form $I(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{M}\mathbf{x} + \mathbf{b}^{\mathrm{T}}\mathbf{x} + c$ (with $\mathbf{M} = \mathbf{M}^{\mathrm{T}} \in \mathbb{R}^{d\times d}$, $\mathbf{b} \in \mathbb{R}^d$, and $c \in \mathbb{R}$).

The next two theorems summarize a few key facts about structure-preserving Runge–Kutta methods.

The next two theorems summarize a few key facts about structure-preserving Runge–Kutta methods.

Theorem

Every Runge–Kutta method preserves linear first integrals.

The next two theorems summarize a few key facts about structure-preserving Runge–Kutta methods.

Theorem

Every Runge–Kutta method preserves linear first integrals.

Proof: $I(\mathbf{x}) = \mathbf{b}^{\mathrm{T}}\mathbf{x} + c$ implies that $\mathbf{D}I(\mathbf{x}) \equiv \mathbf{b}$; thus, by the Lemma, $0 = \mathbf{D}I(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x}) = \mathbf{b} \cdot \mathbf{f}(\mathbf{x})$ for all $\mathbf{x} \in D$.

The next two theorems summarize a few key facts about structure-preserving Runge–Kutta methods.

Theorem

Every Runge–Kutta method preserves linear first integrals.

Proof: $I(\mathbf{x}) = \mathbf{b}^{\mathrm{T}}\mathbf{x} + c$ implies that $\mathbf{D}I(\mathbf{x}) \equiv \mathbf{b}$; thus, by the Lemma, $0 = \mathbf{D}I(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x}) = \mathbf{b} \cdot \mathbf{f}(\mathbf{x})$ for all $\mathbf{x} \in D$.
Consider the R-stage Runge–Kutta method

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h(c_1\mathbf{k}_1 + \cdots + c_R\mathbf{k}_R).$$

The next two theorems summarize a few key facts about structure-preserving Runge–Kutta methods.

Theorem

Every Runge–Kutta method preserves linear first integrals.

Proof: $I(\mathbf{x}) = \mathbf{b}^{\mathrm{T}}\mathbf{x} + c$ implies that $\mathbf{D}I(\mathbf{x}) \equiv \mathbf{b}$; thus, by the Lemma, $0 = \mathbf{D}I(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x}) = \mathbf{b} \cdot \mathbf{f}(\mathbf{x})$ for all $\mathbf{x} \in D$.
Consider the R-stage Runge–Kutta method

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h(c_1\mathbf{k}_1 + \cdots + c_R\mathbf{k}_R).$$

Hence,

$$\begin{aligned} I(\mathbf{y}_{n+1}) - I(\mathbf{y}_n) &= \mathbf{b}^{\mathrm{T}}(\mathbf{y}_{n+1} - \mathbf{y}_n) \\ &= h\,\mathbf{b}^{\mathrm{T}}(c_1\mathbf{k}_1 + \cdots + c_R\mathbf{k}_R) = 0 + \cdots + 0 = 0. \end{aligned}$$

The next two theorems summarize a few key facts about structure-preserving Runge–Kutta methods.

Theorem

Every Runge–Kutta method preserves linear first integrals.

Proof: $I(\mathbf{x}) = \mathbf{b}^{\mathrm{T}}\mathbf{x} + c$ implies that $\mathbf{D}I(\mathbf{x}) \equiv \mathbf{b}$; thus, by the Lemma, $0 = \mathbf{D}I(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x}) = \mathbf{b} \cdot \mathbf{f}(\mathbf{x})$ for all $\mathbf{x} \in D$.
Consider the R-stage Runge–Kutta method

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h(c_1\mathbf{k}_1 + \cdots + c_R\mathbf{k}_R).$$

Hence,

$$\begin{aligned} I(\mathbf{y}_{n+1}) - I(\mathbf{y}_n) &= \mathbf{b}^{\mathrm{T}}(\mathbf{y}_{n+1} - \mathbf{y}_n) \\ &= h\,\mathbf{b}^{\mathrm{T}}(c_1\mathbf{k}_1 + \cdots + c_R\mathbf{k}_R) = 0 + \cdots + 0 = 0. \end{aligned}$$

Theorem

Gauss-collocation methods (i.e. Runge–Kutta methods based on function-evaluations at points of Gaussian quadrature rules) preserve quadratic first integrals.

Simple examples of Gauss-collocation methods are the Gauss–Legendre–Runge–Kutta methods, based on function-evaluations at points of Gauss–Legendre quadrature rules.

Simple examples of Gauss-collocation methods are the Gauss–Legendre–Runge–Kutta methods, based on function-evaluations at points of Gauss–Legendre quadrature rules.

Example

The Gauss–Legendre method of order two is the **implicit midpoint rule**,

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h\,\mathbf{f}\left(t_n + \frac{1}{2}h, \frac{1}{2}\mathbf{y}_n + \frac{1}{2}\mathbf{y}_{n+1}\right)$$

which has Butcher tableau

$$\begin{array}{c|c} 1/2 & 1/2 \\ \hline & 1 \end{array} \quad .$$

Example

The Gauss–Legendre method of order four

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h\left(\frac{1}{2}\mathbf{k}_1 + \frac{1}{2}\mathbf{k}_2\right), \qquad \text{where}$$

$$\mathbf{k}_1 = \mathbf{f}\left(t + \left(\frac{1}{2} - \frac{1}{6}\sqrt{3}\right)h, \mathbf{y}_n + \frac{1}{4}\mathbf{k}_1 + \left(\frac{1}{4} - \frac{1}{6}\sqrt{3}\right)\mathbf{k}_2\right),$$

$$\mathbf{k}_2 = \mathbf{f}\left(t + \left(\frac{1}{2} + \frac{1}{6}\sqrt{3}\right)h, \mathbf{y}_n + \left(\frac{1}{4} + \frac{1}{6}\sqrt{3}\right)\mathbf{k}_1 + \frac{1}{4}\mathbf{k}_2\right)$$

has Butcher tableau

$$\begin{array}{c|cc} \frac{1}{2} - \frac{1}{6}\sqrt{3} & \frac{1}{4} & \frac{1}{4} - \frac{1}{6}\sqrt{3} \\ \frac{1}{2} + \frac{1}{6}\sqrt{3} & \frac{1}{4} + \frac{1}{6}\sqrt{3} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}.$$

Unfortunately, there is no consistent Runge–Kutta method that preserves polynomial first integrals of degree higher than 2; more precisely, the following negative result holds.

Unfortunately, there is no consistent Runge–Kutta method that preserves polynomial first integrals of degree higher than 2; more precisely, the following negative result holds.

Theorem

If $n \geq 3$, then there is no consistent Runge–Kutta method that preserves every polynomial first integral of degree n for every autonomous ODE.

We conclude with a few results concerning the conservation of a structure that is at the heart of classical mechanics: conservation of the symplectic product.

We conclude with a few results concerning the conservation of a structure that is at the heart of classical mechanics: conservation of the symplectic product.

First, we recall the notion of Hamiltonian differential equation from classical mechanics.

We conclude with a few results concerning the conservation of a structure that is at the heart of classical mechanics: conservation of the symplectic product.

First, we recall the notion of Hamiltonian differential equation from classical mechanics.

Definition

A **Hamiltonian differential equation** is an ODE of the form

$$\mathbf{p}' = -\mathbf{D_q} H(\mathbf{p}, \mathbf{q}), \quad \mathbf{q}' = \mathbf{D_p} H(\mathbf{p}, \mathbf{q}), \tag{3}$$

where the function $H : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is called the *Hamiltonian*.

Theorem

The Hamiltonian H is a first integral of (3).

Theorem

The Hamiltonian H is a first integral of (3).

Proof. This follows by applying the chain rule:

$$\frac{\mathrm{d}}{\mathrm{d}t} H(\mathbf{p}, \mathbf{q}) = \mathbf{D_p} H(\mathbf{p}, \mathbf{q}) \cdot \mathbf{p}' + \mathbf{D_q} H(\mathbf{p}, \mathbf{q}) \cdot \mathbf{q}' = 0. \qquad \diamond$$

Theorem

The Hamiltonian H is a first integral of (3).

Proof. This follows by applying the chain rule:

$$\frac{\mathrm{d}}{\mathrm{d}t} H(\mathbf{p}, \mathbf{q}) = \mathbf{D_p} H(\mathbf{p}, \mathbf{q}) \cdot \mathbf{p}' + \mathbf{D_q} H(\mathbf{p}, \mathbf{q}) \cdot \mathbf{q}' = 0. \qquad \diamond$$

Lemma

The ODE (3) *is equivalent to*

$$\mathbf{y}' = \mathbf{J}^{-1} \mathbf{D} H(\mathbf{y}), \tag{4}$$

where $\mathbf{J} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix} \in \mathbb{R}^{2d \times 2d}$ *and* $\mathbf{y} := \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix} \in \mathbb{R}^{2d}$.

The proof of this lemma is left as an exercise.

The following definition is inspired by the previous lemma.

The following definition is inspired by the previous lemma.

Definition

The bilinear map

$$\omega : \mathbb{R}^{2d} \times \mathbb{R}^{2d} \to \mathbb{R}, \quad (\mathbf{a}, \mathbf{b}) \mapsto \omega(\mathbf{a}, \mathbf{b}) \coloneqq \mathbf{a}^{\mathrm{T}} \mathbf{J}\, \mathbf{b}$$

is called the *symplectic product* of $\mathbf{a}$ and $\mathbf{b}$.

Definition

A continuously differentiable map $\mathbf{\Phi} : D \subset \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ is called *symplectic* if

$$\omega(\mathbf{D\Phi}(\mathbf{x})\mathbf{a}, \mathbf{D\Phi}(\mathbf{x})\mathbf{b}) = \omega(\mathbf{a}, \mathbf{b})$$

for every $\mathbf{x} \in D$ and every pair $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^{2d} \times \mathbb{R}^{2d}$.

Definition

A continuously differentiable map $\mathbf{\Phi} : D \subset \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ is called *symplectic* if

$$\omega(\mathbf{D\Phi}(\mathbf{x})\mathbf{a}, \mathbf{D\Phi}(\mathbf{x})\mathbf{b}) = \omega(\mathbf{a}, \mathbf{b})$$

for every $\mathbf{x} \in D$ and every pair $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^{2d} \times \mathbb{R}^{2d}$.

Remark

In other words, a map $\mathbf{\Phi}$ is symplectic if its Jacobian matrix $\mathbf{D\Phi}(\mathbf{x})$ (evaluated at a generic point $\mathbf{x}$) preserves the simplectic product.

Definition

A continuously differentiable map $\mathbf{\Phi} : D \subset \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ is called *symplectic* if

$$\omega(\mathbf{D\Phi}(\mathbf{x})\mathbf{a}, \mathbf{D\Phi}(\mathbf{x})\mathbf{b}) = \omega(\mathbf{a}, \mathbf{b})$$

for every $\mathbf{x} \in D$ and every pair $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^{2d} \times \mathbb{R}^{2d}$.

Remark

In other words, a map $\mathbf{\Phi}$ is symplectic if its Jacobian matrix $\mathbf{D\Phi}(\mathbf{x})$ (evaluated at a generic point $\mathbf{x}$) preserves the simplectic product.

This concept is similar to the property of orthogonal matrices that they preserve the Euclidean inner product, i.e. if $\mathbf{O} \in \mathbb{R}^{d \times d}$ is an orthogonal matrix then $\langle \mathbf{Oa}, \mathbf{Ob} \rangle = \langle \mathbf{a}, \mathbf{b} \rangle$ for every pair $(\mathbf{a}, \mathbf{b})$ in $\mathbb{R}^d \times \mathbb{R}^d$, where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product in $\mathbb{R}^d$.

The following result, due to Poincaré asserts that the

Hamiltonian flow is a symplectic map,

which explains why the concept of symplectic map is so relevant.

The following result, due to Poincaré asserts that the

Hamiltonian flow is a symplectic map,

which explains why the concept of symplectic map is so relevant.

Theorem (Poincaré)

If H is a twice continuously differentiable Hamiltonian, then the flow $\mathbf{\Phi}^t$ of an Hamiltonian differential equation satisfies the following property: for each $\mathbf{x} \in D$ there exists a $\delta > 0$ such that

$$\omega(\mathbf{D\Phi}^t(\mathbf{x})\mathbf{a}, \mathbf{D\Phi}^t(\mathbf{x})\mathbf{b}) = \omega(\mathbf{a}, \mathbf{b})$$

for every $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^{2d} \times \mathbb{R}^{2d}$ and all $t \in [0, \delta)$.

Since Hamiltonian flows are symplectic, we are interested in symplectic one-step methods, in the following sense.

Since Hamiltonian flows are symplectic, we are interested in symplectic one-step methods, in the following sense.

Definition

Consider (4) subject to the initial condition $\mathbf{y}(0) = \mathbf{x}$, for $\mathbf{x} \in D$, and let $\mathbf{x} \mapsto \boldsymbol{\Psi}(0, \mathbf{x}; h)$ be a one-step method for (4), which maps the initial datum $\mathbf{x} \in D$ into a numerical approximation $\boldsymbol{\Psi}(0, \mathbf{x}; h) \in \mathbb{R}^d$ of $\mathbf{y}(h) \in \mathbb{R}^d$ over a single time step $h > 0$.

Since Hamiltonian flows are symplectic, we are interested in symplectic one-step methods, in the following sense.

Definition

Consider (4) subject to the initial condition $\mathbf{y}(0) = \mathbf{x}$, for $\mathbf{x} \in D$, and let $\mathbf{x} \mapsto \boldsymbol{\Psi}(0, \mathbf{x}; h)$ be a one-step method for (4), which maps the initial datum $\mathbf{x} \in D$ into a numerical approximation $\boldsymbol{\Psi}(0, \mathbf{x}; h) \in \mathbb{R}^d$ of $\mathbf{y}(h) \in \mathbb{R}^d$ over a single time step $h > 0$.

The one-step method $\boldsymbol{\Phi}_h : \mathbf{x} \mapsto \boldsymbol{\Psi}(0, \mathbf{x}; h)$ is said to be *symplectic* if $\boldsymbol{\Phi}_h : \mathbf{x} \mapsto \boldsymbol{\Psi}(0, \mathbf{x}; h)$ defines a symplectic map on every compact subset $K \subset D$, whenever H is twice continuously differentiable and $h > 0$ is sufficiently small.

The following theorem provides a convenient sufficient condition for a Runge–Kutta being symplectic, although for an arbitrary one-step method one would still need to appeal to the previous definition to verify that the method in question is symplectic.

The following theorem provides a convenient sufficient condition for a Runge–Kutta being symplectic, although for an arbitrary one-step method one would still need to appeal to the previous definition to verify that the method in question is symplectic.

Theorem

Every Runge–Kutta method that preserves quadratic first integrals is symplectic.