Numerical integration of ODEs

Cauchy problem

$$\begin{cases} \int_{t_n}^{t_n+h} \dot{\vec{y}}(t)\, dt = \int_{t_n}^{t_n+h} \vec{f}(\vec{y}(t), t)\, dt \\ \vec{y}(0) = \vec{y}_0 \end{cases}$$

h

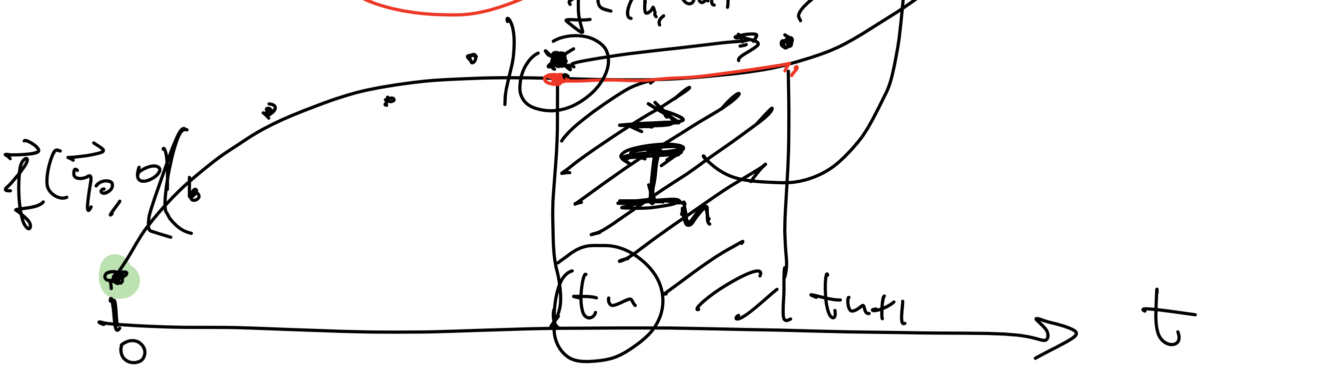
$0 \quad h \qquad\qquad t_n$

$t_0 \quad t_1 \qquad\qquad nh$

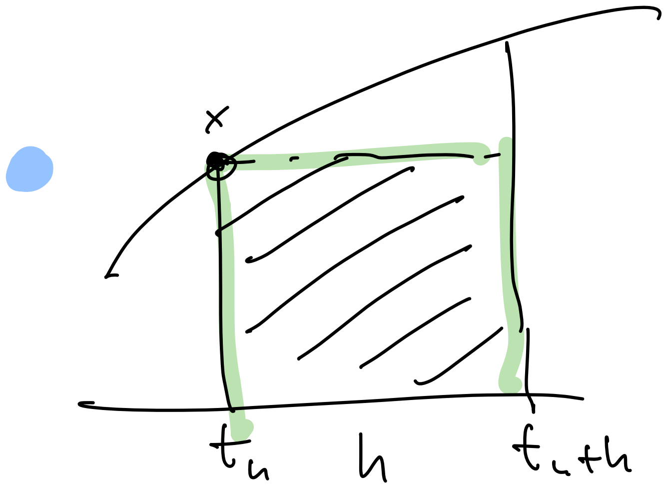
$$\vec{y}_{n+1} \simeq \vec{y}(t_{n+1})$$

$$= \vec{\mathcal{Y}}_{(n)}\left((\vec{y}_{n+1}), \vec{y}_n, \dots, \vec{y}_0 ; \vec{f}, h\right)$$

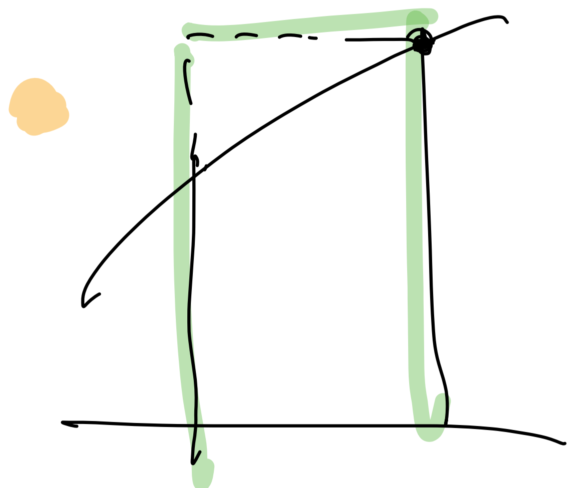
$$\vec{y}_n \rightarrow \vec{y}_{n+1}$$



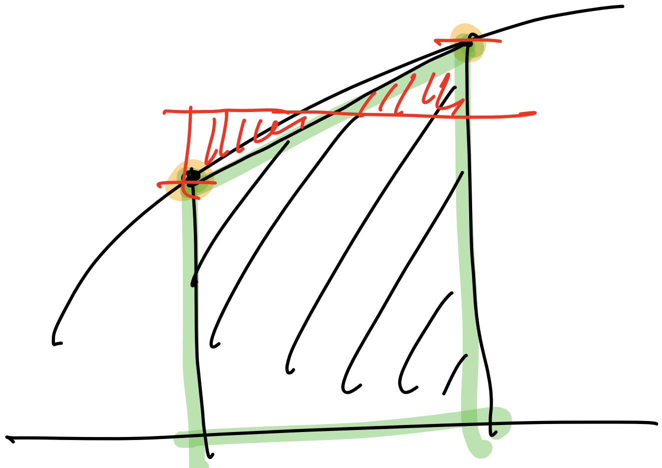
$$\vec{y}(t_n) \to \vec{y}(t_{n+1}) = \vec{y}(t_n) + \underbrace{\int_{t_n}^{t_n+h} dt\, \vec{f}(\vec{y}(t), t)}_{\vec{I}_n}$$



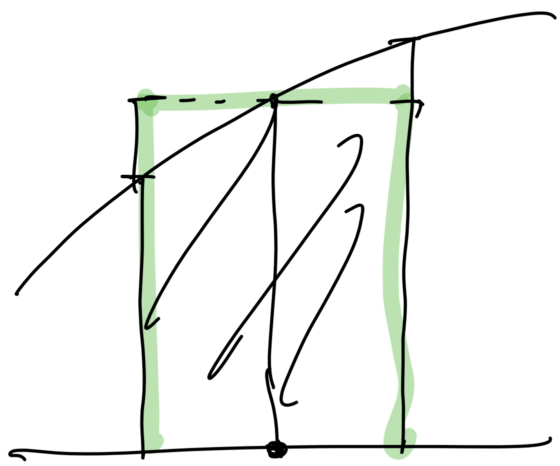
(forward Euler)



(backward Euler)



trapezoidal scheme



midpoint scheme

$$\vec{y}(t_{n+1}) = \vec{y}(t_n) + \int_{t_n}^{t_n+h} dt\, \underbrace{\vec{f}(\vec{y}(t), t)}$$

$$\simeq \vec q(t_n) + \int_{t_n}^{t_n+h} dt \left[\vec f(\vec q(t_n), t_n) + (t-t_n) \frac{d}{dt}\vec f(\vec q, t)\Big|_{t=t_n} + \dots \right]$$

$$= \underline{\vec q(t_n) + h\, \vec f(\vec q(t_n), t_n)} + \underbrace{\cancel{\frac{h^2}{2} \frac{d}{dt} \vec f\Big|_{t_n}} + O(h^3)}_{O(h^2)}$$

$$= \vec q(t_n) + h\, \vec f(\vec q(t_{n+1}), t_{n+1}) + O(h^2)$$

Euler method

$$\vec q_0 = \vec q(0)$$

$$\vec q_1 = h\, \vec f(\vec q(0), 0) + \vec q_0$$

...

$$\boxed{\vec q_{n+1} = \vec q_n + h\, \vec f(\vec q_n, t_n)}$$

Forward Euler method

$\vec q_n \simeq \vec q(t_n)$

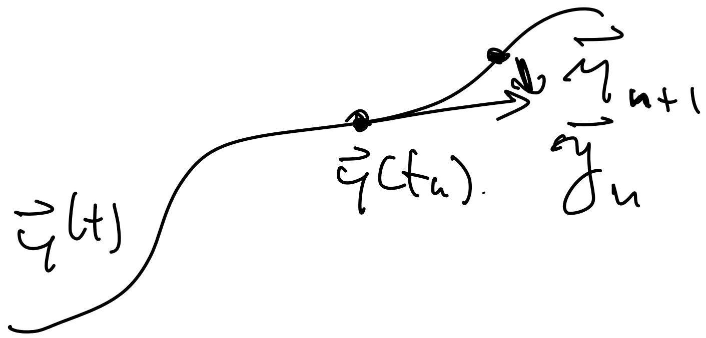
num. ex.

how good is the approximation

local error for a method

$$\vec{y}_{n+1} = \vec{\mathcal{Y}}_n \big(\underbrace{(\vec{y}_{n+1}), \vec{y}_n, \dots, \vec{y}_0; \vec{f}, h}\big)$$

$$\vec{\eta}_{n+1} = \vec{y}(t_{n+1}) - \vec{\mathcal{Y}}_n\big((\vec{y}(t_{n+1})), \vec{y}(t_n), \dots, \vec{y}_0; \vec{f}, h\big)$$



Method of order p

$$|\vec{\eta}_{n+1}| \simeq O(h^{p+1})$$

Consistency of order p of a method

$[0, T]$

$$\max_{n \in [0, \frac{T}{h}]} |\vec{\eta}_n| = h\,\theta(h)$$

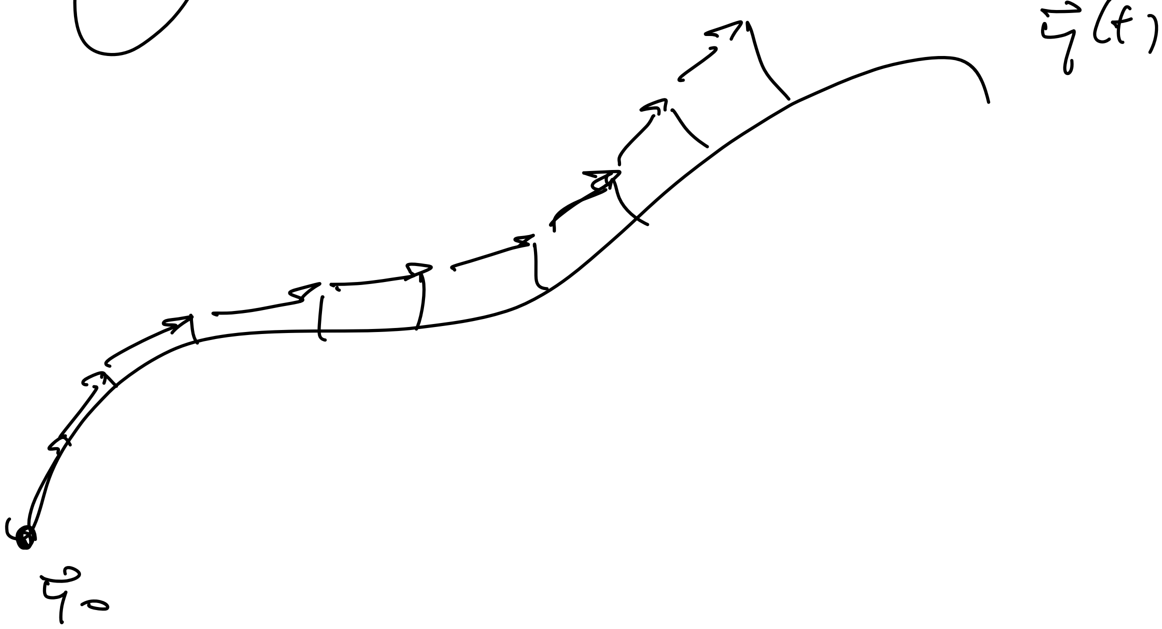
$$\theta(h) \xrightarrow[h \to 0]{} h^p$$

Convergence of order p of a method

global error $[0, T]$

$$\vec{e}_n = \vec{y}(t_n) - \vec{y}_n$$

$$\max_{n \in [0, \frac{T}{h}]} |\vec{e}_n| \underset{h \to 0}{\sim} O(h^p)$$

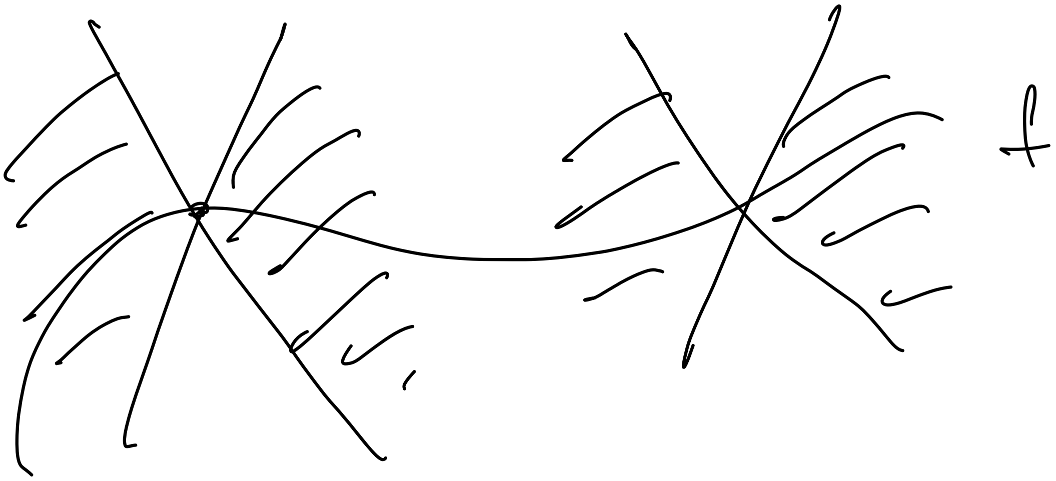


General idea : convergence $\simeq$ consistency

For instance : if you assume that $\vec{f}$ is Lipschitz - continuous

$$\| \vec{f}(\vec{x}, t) - \vec{f}(\vec{y}, t) \| \leq L \| \vec{x} - \vec{y} \|$$

$$\forall t \in [0, T) \quad \forall \vec{x}, \vec{y} \in A$$



$\Rightarrow$ Euler method consistent of order 1 is also convergent of order 1

Consequence : $M = \frac{\tau}{h}$ steps

accuracy $O(h^2) \times \frac{\tau}{h} \sim O(h)$

$\epsilon = \| \vec{q}(\tau) - \vec{q}_M \| \sim O(h)$

ϵ accuracy $\sim h$

$T \sim M \times N \sim N \frac{\tau}{h} \sim \frac{N}{\epsilon} \tau$

time to solution

dimensions of $\vec{q}$

EULER

Refinements to the Euler method

$$\epsilon \sim h^p \qquad p > 1$$

$$T \sim \frac{N}{\epsilon^{1/p}} \, \tau$$

→ better scalings in ϵ

→ "stiff" ODEs

Examples

backward Euler:

$$\vec{y}_{n+1} = \vec{y}_n + h \vec{f}(\vec{y}_{n+1}, t_{n+1})$$

trapezoidal scheme

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

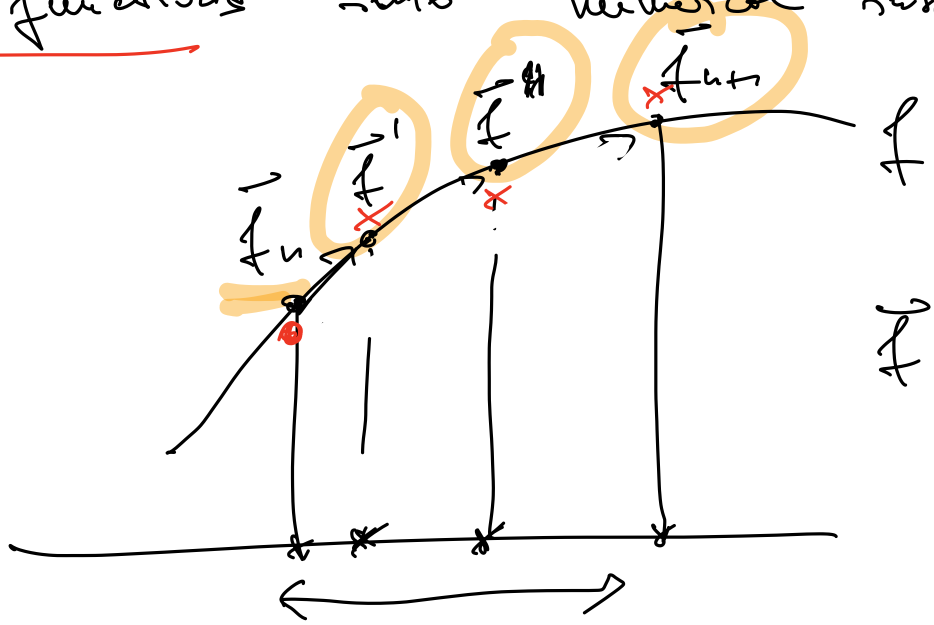
$$\vec{y}(t_{n+1}) = \vec{y}(t_n) + \frac{h}{2} \left[\underline{\vec{f}(\vec{y}(t_n), t_n)} + \underbrace{\vec{f}(\vec{y}(t_{n+1}), t_{n+1})}_{\text{Taylor expand around } t_n} \right]$$

$$= \vec{q}(t_n) + h \underbrace{\vec{f}(\vec{q}(t_n), t_n)}_{\dot{\vec{q}}(t_n)} + \frac{1}{2}h^2 \underbrace{\left.\frac{d\vec{f}}{dt}\right|_{t_n}}_{\ddot{\vec{q}}(t_n)} + o(h^3)$$

$$= \vec{q}(t_n) + h\,\dot{\vec{q}}(t_n) + \frac{1}{2}h^2\,\ddot{\vec{q}}(t_n) + o(h^3)$$

consistency of order 2.

$\boxed{\text{Translate}}$ "any" scheme of numerical integration of functions into numerical integration of ODEs



$\vec{f}(\vec{q}(t), t)$

Runge-Kutta methods

m-stage RK method

$\vec{k}_1, \quad \vec{k}_2, \quad \ldots, \quad \vec{k}_{(m)}$ RK vectors

22 22 22

$\vec{f}^{\,1} - \vec{y}_n, \quad \vec{f}^{\,11} - \vec{y}_n, \quad \vec{f}^{\,(m)} - \vec{y}_n,$

more generally

$\vec{y}_n + a_{11}\vec{k}_1 + a_{12}\vec{k}_2 + \ldots$

general scheme (explicit)

forward Euler

$$\vec{k}_1 = h\,\vec{f}(\vec{y}_n, t_n) \quad \rightarrow \quad \vec{y}_{n+1} = \vec{y}_n + \vec{k}_1$$

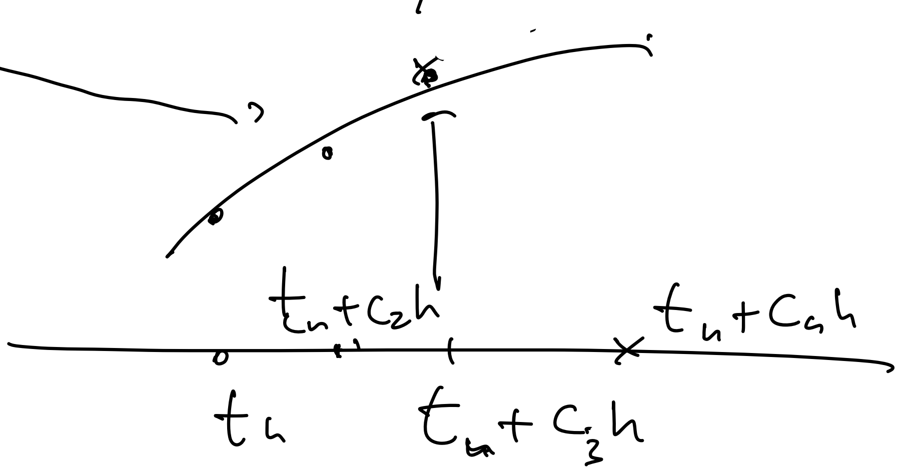
$$\vec{k}_2 = h\,\vec{f}(\vec{y}_n + a_{21}\vec{k}_1,\ t_n + c_2 h)$$

ex. $a_{21} = 1$, $c_2 = 1$

$$\vec{k}_2 = h\,\vec{f}(\overbrace{\vec{y}_n + \vec{k}_1}^{\vec{y}_{n+1}},\ t_n + h)$$

backward Euler

$$\vec{k}_3 = h\,\vec{f}(\vec{y}_n + a_{32}\vec{k}_2 + a_{31}\vec{k}_1,\ t_n + c_3 h)$$



...

$$\vec{k}_m = h\vec{f}\left(\vec{y}_n + a_{m\,m-1}\vec{k}_{m-1} + \dots + a_{m1}\vec{k}_1,\; t_n + c_m h\right)$$

$$\vec{y}_{n+1} = \vec{y}_n + h\sum_{i=1}^{m} b_i \vec{k}_i$$

$$\{a_{ij}\},\ \{b_i\},\ \{c_i\}$$

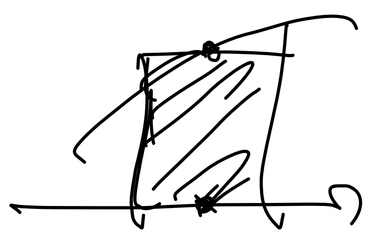
Butcher's tableau

$$\begin{array}{c|cccc}
c_1 & a_{11} & a_{12} & \dots & a_{1m} \\
c_2 & \vdots & & & \\
\vdots & \vdots & & & \\
c_m & a_{m1} & a_{m2} & \dots & a_{mm} \\
\hline
 & b_1 & b_2 & \dots & b_m
\end{array}$$

only non-zero elements for explicit RK methods

Examples

→ explicit midpoint method



$\to \quad \vec{k}_1 = h\,\vec{f}(\vec{y}_n, t_n)$

$\vec{k}_2 = h\,\vec{f}\left(\vec{y}_n + \frac{\vec{k}_1}{2},\ t_n + \frac{h}{2}\right) \quad \leftarrow$

approx. to the mid point

$\vec{y}_{n+1} = \vec{y}_n + \vec{k}_2$

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \hline & 0 & 1 \end{array}$$

$\to$ RK(4)   4-stage method

Related numerical integration scheme

Simpson's $\frac{1}{3}$ rule

$$\int_a^b \vec{f}(t)\,dt = \frac{(b-a)}{6}\left[\vec{f}(a) + \vec{f}(b) + 4\vec{f}\left(\frac{a+b}{2}\right)\right] + O\underbrace{(b-a)}_{h}^{5}$$

$$\vec{k}_1 = h\,\vec{f}(\vec{y}_n, t_n) \qquad \left(\simeq \vec{f}(a)\right)$$

$$\vec{k}_2 = h\,\vec{f}\left(\vec{y}_n + \frac{\vec{k}_1}{2}, t_n + \frac{h}{2}\right) \qquad \simeq \vec{f}\left(\frac{a+b}{2}\right)$$

$$\vec{k}_3 = h\,\vec{f}\left(\vec{y}_n + \frac{\vec{k}_2}{2}, t_n + \frac{h}{2}\right) \qquad \simeq \vec{f}\left(\frac{a+b}{2}\right)$$

$$\vec{k}_4 = h\,\vec{f}(\vec{y}_n + \vec{k}_3, t_n + h) \qquad \simeq \vec{f}(b)$$

$$\vec{y}_{n+1} = \vec{y}_n + \frac{1}{6}\left[\vec{k}_1 + 2\vec{k}_2 + 2\vec{k}_3 + \vec{k}_4\right]$$

$$\begin{array}{c|cccc} 0 & 0 & & & \\ \frac{1}{2} & \frac{1}{2} & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ 1 & 0 & 0 & 1 & 0 \\ \hline & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{array}$$

Method of order 4 (in consistency)

for m stages

RK methods with $m > 4$ stages are of order $p < m$

Stability of a Cauchy problem

$$\begin{cases} \dot{\vec{q}}(t) = \vec{f}(q(t), t) \\ \vec{q}(0) = \vec{q}_0 \end{cases}$$

$$\begin{cases} \dot{\vec{z}}(t) = \vec{f}(\vec{z}(t), t) + \vec{\delta}(t) \\ \vec{z}(0) = \vec{q}_0 + \vec{\delta}_0 \end{cases}$$

problem stable $[0, \tau]$ if $\|\vec{\delta}(t)\| < \varepsilon$, $t \in [0, \tau]$

$\Rightarrow \|\vec{z}(t) - \vec{q}(t)\| < C\varepsilon$ $\exists\, C > 0$

Restrict to stable Cauchy problems

Zero-stability of a numerical method

More general formulation of a numerical method

$\vec{q}_n + h\vec{f}(\qquad)$ method

$$\vec{q}_{n+1} = \vec{\mathcal{Y}}_n\left((\vec{q}_{n+1}), \ldots q_0; \vec{f}, h\right)$$

$\forall$ more general

$$\sum_{j=0}^{k} \alpha_j \vec{q}_{n+j} = h\,\vec{\Phi}_{\vec{f}}(\vec{q}_{n+k}, \ldots, \vec{y}_0; t_n, h)$$

Vectors $\vec{q}_n$ appear only as arguments of $\vec{f}$

RK

$$\vec{y}_{n+1} - \vec{y}_n = h\,\vec{\Phi}(\qquad)$$

$$\begin{cases} \alpha_0 = -1 \\ \alpha_1 = 1 \end{cases}$$

$\vec{q}_0$ initial condition

Solution $\{\vec{q}_n\}$ with the without $\vec{\Phi}$ with $\vec{q}_0$

solution $\{\vec{z}_n\}$

$$\sum_{j=0}^{k} \alpha_j \vec{z}_{n+j} = h\left[\vec{\Phi}_f(\qquad) + \vec{\delta}_{n+k}\right]$$

$$\vec{z}_0 = \vec{q}_0 + \vec{\delta}_0$$

Method is zero-stable if

$\forall n \in [0, \frac{T}{h}]$ $\Rightarrow$ $\exists S > 0$:

$\|\vec{\delta}_n\| < \varepsilon$ $\quad \|\vec{z}_n - \vec{y}_n\| < S\varepsilon$

Fundamental theorem of numerical analysis

$\boxed{\begin{array}{l} \cdot \text{ consistency of order } p \\ \cdot \text{ zero-stability} \end{array}} \Leftrightarrow \boxed{\begin{array}{l} \text{convergence} \\ \text{of order } p \end{array}}$

Proof of zero-stability

characteristic polynomial

$$p(x) = \sum_j \alpha_j x^j$$

RK $p(x) = x - 1$

$\boxed{\begin{array}{l} \text{Method is} \\ \text{zero-stable} \end{array}} \Leftrightarrow \boxed{\begin{array}{l} p(x) \text{ has roots} \\ |x_r| \leq 1, \text{ and} \\ |x_r| = 1 \text{ are simple} \end{array}}$

$\Rightarrow$ RK methods are convergent of order p