



# VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

## FAKULTA STROJNÍHO INŽENÝRSTVÍ

FACULTY OF MECHANICAL ENGINEERING

## ÚSTAV MATEMATIKY

INSTITUTE OF MATHEMATICS

# JEDNOKROKOVÉ NUMERICKÉ METODY PRO ŘEŠENÍ ZPOŽDĚNÝCH DIFERENCIÁLNÍCH ROVNIC

ONE-STEP NUMERICAL METHODS FOR DELAY DIFFERENTIAL EQUATIONS

## BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

## AUTOR PRÁCE

AUTHOR

Samuel Štěpán

## VEDOUCÍ PRÁCE

SUPERVISOR

doc. Ing. Petr Tomášek, Ph.D.

BRNO 2022



# Zadání bakalářské práce

Ústav: Ústav matematiky  
Student: **Samuel Štěpán**  
Studijní program: Matematické inženýrství  
Studijní obor: bez specializace  
Vedoucí práce: **doc. Ing. Petr Tomášek, Ph.D.**  
Akademický rok: 2021/22

Ředitel ústavu Vám v souladu se zákonem č.111/1998 o vysokých školách a se Studijním a zkušebním řádem VUT v Brně určuje následující téma bakalářské práce:

## Jednokrokové numerické metody pro řešení zpožděných diferenciálních rovnic

### Stručná charakteristika problematiky úkolu:

Zpožděné diferenciální rovnice vykazují řadu zajímavých vlastností ve srovnání s, v jistém smyslu, odpovídajícími obyčejnými diferenciálními rovnicemi [1]. Rovněž numerické řešení vyžaduje speciální přístup již vzhledem k jiné struktuře počátečního problému. Byla odvozena řada numerických metod určených speciálně pro zpožděné diferenciální rovnice, a to i s ohledem na typ zpožděného argumentu.

### Cíle bakalářské práce:

Nastudování jedнокrokových metod pro řešení počátečních problémů diferenciálních rovnic s konstantním zpožděním.

Realizace vybraných numerických schémat v prostředí MATLAB.

Studium vlastností numerického řešení na vybraných úlohách.

### Seznam doporučené literatury:

BAKER, Christopher T. H. Retarded differential equations. Journal of computational and applied mathematics [online]. AMSTERDAM: Elsevier B.V, 2000, 125(1), 309-335 [cit. 2021-10-22]. ISSN 0377-0427. Dostupné z: doi:10.1016/S0377-0427(00)00476-3. [1]

BELLEN, Alfredo, ZENNARO, Marino. Numerical methods for delay differential equations. Oxford: Oxford University Press, 2005, xiv, 395 s. ISBN 0-19-850654-6.

JIAOXUN, Kuang, YUHAO, Cong. Stability of numerical methods for delay differential equations. Beijing: Science Press, 2005, vii, 295 s. ISBN 978-7-03-016317-2.

Termín odevzdání bakalářské práce je stanoven časovým plánem akademického roku 2021/22

V Brně, dne

L. S.

---

prof. RNDr. Josef Šlapal, CSc.  
ředitel ústavu

---

doc. Ing. Jaroslav Katolický, Ph.D.  
děkan fakulty



## **Abstrakt**

Bakalářská práce se zabývá formulací a vlastnostmi jednokrokových numerických metod pro počáteční problémy diferenciálních rovnic s konstantním zpožděním. Zejména jsou v této souvislosti diskutovány spojité Rungeovy-Kuttovy metody. Vybrané numerické algoritmy jsou implementovány ve vývojovém prostředí MATLAB. Ty jsou pak na vybraných úlohách porovnány s řešičem dde23.

## **Summary**

The bachelor thesis deals with the formulation and properties of one-step numerical methods for the initial problems of differential equations with constant delay. In particular, continuous Runge-Kutta methods are discussed in this context. Some numerical algorithms are implemented in the MATLAB environment. These are then compared with the dde23 solver on specific problems.

## **Klíčová slova**

Diferenciální rovnice se zpožděním, konstantní zpoždění, spojitá Rungeova-Kuttova metoda, variabilní krok.

## **Keywords**

Delay differential equation, constant delay, continuous Runge-Kutta method, variable stepsize.

ŠTĚPÁN, S. *Jednokrokové numerické metody pro řešení zpožděných diferenciálních rovnic*. Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2022. 59 s. Vedoucí doc. Ing. Petr Tomášek, Ph.D.



## Prohlášení

Prohlašuji, že svou bakalářskou práci na téma “Jednokrokové numerické metody pro řešení zpožděných diferenciálních rovnic“ jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury.

Jako autor uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení §11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 SB.





## Poděkování

Rád bych poděkoval vedoucímu bakalářské práce panu doc. Ing. Petrovi Tomáškoví, Ph.D. za odborné vedení, konzultace a trpělivost.



# Obsah

Úvod	1
<b>1 Diferenciální rovnice s konstantním zpožděním</b>	<b>2</b>
1.1 Primární a sekundární body nespojitosti	2
1.2 Existence a jednoznačnost	4
<b>2 Obyčejné diferenciální rovnice - spojitě numerické metody</b>	<b>6</b>
<b>3 Zpožděné diferenciální rovnice - numerické řešení</b>	<b>9</b>
3.1 Feldsteinova metoda	9
3.2 Bellmanova metoda kroků	10
3.3 Standardní přístup pomocí spojitých ODR metod	10
<b>4 Spojitě Rungeovy-Kuttovy metody pro obyčejné diferenciální rovnice</b>	<b>13</b>
4.1 Spojitě prodloužení RK metod	14
4.1.1 Interpolant první třídy	14
4.1.2 Interpolant druhé třídy	14
4.2 Analýza interpolantu první třídy	16
4.2.1 Kolokační metody	17
4.3 Aplikace odhadu lokální chyby	22
<b>5 Rungeovy-Kuttovy metody pro zpožděné diferenciální rovnice</b>	<b>24</b>
<b>6 Odhad lokální chyby a variabilní krok</b>	<b>27</b>
6.1 Řád hlavní a pomocné metody	27
6.1.1 Pomocná metoda řádu $p' = p - 1$	29
6.1.2 Pomocná metoda řádu $p' = p + 1$	29
6.1.3 Spojitá pomocná metoda	30
6.2 Šíření chyby a kontrola délky kroku	30
6.2.1 Stejnoměrný řád $q = p$	33
6.2.2 Stejnoměrný lokální řád $q = p - 1$	33
6.2.3 Shrnutí	33
6.3 Implementace	34
<b>7 Analýza stability ZDR</b>	<b>38</b>
7.1 Obecný nelineární systém ZDR s konstantním zpožděním	38
7.2 Lineární skalární testovací rovnice	38
<b>8 Příklady</b>	<b>41</b>
<b>Závěr</b>	<b>52</b>
<b>Seznam zkratk</b>	<b>57</b>
<b>Seznam příloh</b>	<b>59</b>



# Úvod

V této bakalářské práci se budeme zabývat numerickými metodami pro *zpožděné diferenciální rovnice* (ZDR) s konstantním zpožděním. Z teorie *obyčejných diferenciálních rovnic* (ODR) je známo, že každou ODR řádu  $n > 1$  lze převést na systém ODR řádu 1. To platí i u ZDR, proto se v této práci budeme zabývat systémem zpožděných diferenciálních rovnic řádu 1 s počáteční podmínkou:

$$\begin{cases} y'(t) = f(t, y(t), y(t - \tau)), & t_0 \leq t \leq t_f, \\ y(t) = \phi(t), & t \leq t_0, \end{cases} \quad (0.0.1)$$

kde  $f : [t_0, t_f] \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $\tau$  je konstantní zpoždění a  $\phi : [t_0 - \tau, t_0] \rightarrow \mathbb{R}^d$  je počáteční funkce.

Vzhledem k přítomnosti zpožděného argumentu  $\alpha(t) = t - \tau$  vykazuje řešení ZDR některé zajímavé vlastnosti, které se mohou jevit jako nečekané v souvislosti se znalostí analogií v rámci ODR. U numerického řešení se tyto specifické vlastnosti přímo promítají do konstrukcí numerických schémat a algoritmů.

V první kapitole ilustrujeme vznik nespojitostí z důvodu přítomnosti zpožděného argumentu  $\alpha(t) = t - \tau$  a uvedeme podmínky pro existenci a jednoznačnosti řešení. Protože numerické řešení ZDR může vyžadovat interpolaci, v druhé kapitole se budeme věnovat spojitým numerickým metodám pro řešení ODR. Díky nim budeme schopni zavést v třetí kapitole základní metodu pro řešení ZDR, a to tzv. *metodu kroků*. Ve čtvrté kapitole se budeme do hloubky zabývat spojitými Rungeovými-Kuttovými metodami pro ODR, abychom je v páté kapitole mohli implementovat pro ZDR. Efektivní numerická metoda pracuje s mechanismem kontroly délky kroku, proto se v šesté kapitole budeme věnovat odhadu lokální chyby a implementaci metody s variabilním krokem. Kapitola 7 uvádí stručnou analýzu stability pro obecný nelineární systém ZDR a lineární systém ZDR s reálnými koeficienty. V kapitole 8 ilustrujeme použití vybraných numerických metod na vybraných příkladech. Jednotlivé algoritmy realizované v Matlabu lze nalézt v příloze práce.

# 1 Diferenciální rovnice s konstantním zpožděním

Budeme se zabývat především řešením počáteční úlohy (0.0.1). Nejdříve ilustrujme na následujícím příkladu jednu ze specifických vlastností ZDR.

## Příklad 1.0.1

Uvažujme ZDR s počáteční podmínkou:

$$\begin{cases} y'(t) = -y(t-1), & t \geq 0, \\ y(t) = 1, & t \leq 0. \end{cases} \quad (1.0.1)$$

Řešení je vykresleno v Obrázku 1.1. Protože  $y'(0)^- = 0$  a  $y'(0)^+ = -y(-1) = -1$ , derivace funkce  $y'(t)$  má skok v bodě  $t = 0$ . Druhá derivace  $y''(t)$  je tvaru:

$$y''(t) = -y'(t-1),$$

a proto má skok v bodě  $t = 1$ . Třetí derivace  $y'''(t)$  je tvaru:

$$y'''(t) = -y''(t-1) = y'(t-2),$$

a proto má skok v bodě  $t = 2$ . Takto lze pokračovat pro násobky vyšší derivace, kdy skok se projeví v odpovídajícím násobku zpoždění, tj.  $y^{(n)}(t)$  má nespojitost typu skoku v bodě  $t = n - 1$ .

Ilustrovali jsme tedy, že zpožděný argument  $t - \tau$  způsobuje skokové nespojitosti v  $y'$  nebo ve vyšších derivacích  $y^{(n)}$ ,  $n > 1$ . Je známo, že každá jednokroková numerická metoda na řešení počátečního problému dosahuje určitého řádu přesnosti, za předpokladu, že řešení je dostatečně hladké na každém intervalu  $[t_n, t_{n+1}]$ . Přesněji, pro metodu řádu  $p$  obvykle požadujeme, aby řešení bylo spojitě alespoň třídy  $C^{p+1}$  na intervalu  $[t_n, t_{n+1}]$ . Proto je potřeba, abychom do sítě (používané numerickou metodou) zahrnuli všechny body nespojitosti  $y^{(s)}$ , kde  $s$  je alespoň  $s = 0, 1, \dots, p + 1$ .

## 1.1 Primární a sekundární body nespojitosti

Uvažujme počáteční problém pro ZDR (0.0.1), kde  $d = 1$ . Předpokládejme, že řešení  $y(t)$  nenavazuje hladce na počáteční funkci  $\phi(t)$ , tedy v  $t_0$  platí  $\phi'(t_0)^- \neq y'(t_0)^+ = f(t_0, \phi(t_0), \phi(t_0 - \tau))$ . Pokud funkce  $f$  a  $\phi$  jsou spojitě, potom i  $y'(t)$  je spojitá pro  $t > t_0$ . Dále pokud  $f$  a  $\phi$  jsou diferencovatelné, potom  $y''(t)$  existuje pro všechna  $t$  kromě bodu  $\xi_1$ :

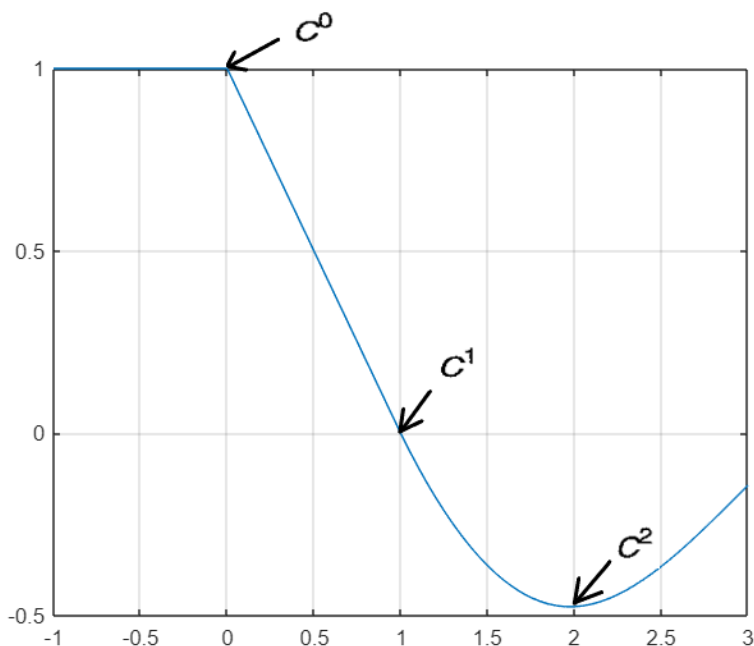
$$\xi_1 = t_0 + \tau.$$

Pro jakoukoliv hladkou funkci  $f(t, y, x)$  můžeme napsat:

$$\begin{aligned} y''(t)^\pm &= \frac{\partial f}{\partial t}(t, y(t), y(t-\tau)) + \frac{\partial f}{\partial y}(t, y(t), y(t-\tau)) y'(t) \\ &\quad + \frac{\partial f}{\partial x}(t, y(t), y(t-\tau)) y'(t-\tau)^\pm \end{aligned} \quad (1.1.1)$$

a tedy

$$\begin{aligned} y''(t)^+ &= \frac{\partial f}{\partial t}(\xi_1, y(\xi_1), y(t_0)) + \frac{\partial f}{\partial y}(\xi_1, y(\xi_1), y(t_0)) y'(\xi_1) \\ &\quad + \frac{\partial f}{\partial x}(\xi_1, y(\xi_1), y(t_0)) y'(t_0)^+ \end{aligned} \quad (1.1.2)$$



Obrázek 1.1: Řešení počátečního problému (1.0.1)

a

$$y''(t)^- = \frac{\partial f}{\partial t}(\xi_1, y(\xi_1), y(t_0)) + \frac{\partial f}{\partial y}(\xi_1, y(\xi_1), y(t_0))y'(\xi_1) + \frac{\partial f}{\partial x}(\xi_1, y(\xi_1), y(t_0))y'(t_0)^-. \quad (1.1.3)$$

Protože předpokládáme, že  $\phi'(t_0)^-$  je různé od  $y'(t_0)^+$ ,  $y''$  neexistuje v bodě  $\xi_1$  a její prodloužení ( $y''(\xi_1) = y''(\xi_1)^+$ ) má v  $\xi_1$  skokovou nespojitost.

Tyto skoky pro  $y''$  budeme nazývat *primární nespojitost 1-třídy*. Pomocí diferencování (1.1.1) můžeme pozorovat, že bod primární nespojitosti první třídy  $\xi_1$  způsobí vznik *primární nespojitosti 2-třídy* v  $y'''$  v bodě  $\xi_2 = \xi_1 + \tau$ .

Obecně tedy bod *primární nespojitosti k-třídy*  $\xi_k$  způsobí vznik *primární nespojitosti (k + 1)-třídy* v  $y^{(k+2)}$  v bodě  $\xi_{k+1} = t_0 + (k + 1)\tau$ , takže řešení (0.0.1) se stává hladší, jak se zvyšuje třída primární nespojitosti. Zlepšení regularity  $y(t)$  bude označováno jako *vyhlazování řešení*.

**Věta 1.1.1** *Pokud  $\xi_i$  je bod primární nespojitosti a v bodě  $\xi_i$  má funkce  $y(t)$  spojitě derivace až do řádu  $\omega - 1$ , potom je  $y(t)$  spojitě diferencovatelná v bodě  $\xi_{i+1} = t_0 + (i + 1)\tau$ .*

Tento jev budeme nazývat *obecným vyhlazováním* a s důkazem této věty se odkazujeme na Nevese a Feldsteina [28].

**Definice 1.1.2** *Bod nespojitosti  $\xi$  je řádu  $k$ , pokud  $y^{(s)}(\xi)$  existuje pro  $s = 0, \dots, k$  a  $y^{(k)}$  je Lipschitzovsky spojitá v  $\xi$ .*

Můžeme poznamenat, že každý bod nespojitosti třídy  $p$  má řád  $k \geq p$ .



Funkce  $f$  a  $\phi$  nemusí být spojité v jejich derivacích vzhledem k  $t$ . Tyto nespojitosti se také budou šířit stejným způsobem díky zpožděnému argumentu  $t - \tau$  pomocí stejných pravidel jako primární nespojitosti. Tyto nespojitosti budeme nazývat *sekundární nespojitosti*.

Stejně jako u primárních nespojitostí, body sekundární nespojitosti musí být také zahrnuty v síti za účelem zachování přesnosti daného řádu.

## 1.2 Existence a jednoznačnost

**Věta 1.2.1 (Lokální existence)** *Uvažujme počáteční problém pro diferenciální rovnici s konstantním zpožděním:*

$$\begin{cases} y'(t) = f(t, y(t), y(t - \tau)), & t \geq t_0, \\ y = \phi(t), & t \leq t_0. \end{cases} \quad (1.2.1)$$

*Nechť  $U \subseteq \mathbb{R}^d$  je okolí  $\phi(t_0)$  a  $V \subseteq \mathbb{R}^d$  je okolí  $\phi(t_0 - \tau)$  a předpokládejme, že funkce  $f(t, u, v)$  je spojitá vzhledem k  $t$  a Lipschitzovsky spojitá vzhledem k  $u$  a  $v$  na množině  $[t_0, t_0 + h] \times U \times V$  pro nějaké  $h > 0$ . Dále předpokládejme, že počáteční funkce  $\phi(t)$  je Lipschitzovsky spojitá pro  $t \leq t_0$ . Potom problém (1.2.1) má jednoznačné řešení v  $[t_0, t_0 + \delta)$  pro nějaké  $\delta > 0$  a řešení závisí spojitě na počáteční funkci.*

Důkaz nalezneme v [6].

Následující výsledky lze získat jako konkrétní případy obecnějších vět dokázaných Halem [19] a El'sgol'tsem a Norkinem [7].

**Věta 1.2.2 (Globální existence)** *Pokud jednoznačné maximální řešení (1.2.1) je omezené, tak podle Věty 1.2.1 existuje řešení na celém intervalu  $[t_0, t_f]$ .*

Za účelem aplikace věty o globálním řádu, potřebujeme prvně omezení pro řešení. To bude dáno následujícím důsledkem:

**Důsledek 1.2.3** *Předpokládejme navíc, že funkce  $f(t, u, v)$  splňuje podmínku:*

$$\| f(t, u, v) \| \leq M(t) + N(t) (\| u \| + \| v \|) \quad (1.2.2)$$

*na intervalu  $A \subseteq \langle t_0, t_f \rangle \times \mathbb{R}^d \times \mathbb{R}^d$ , kde  $M(t)$  a  $N(t)$  jsou spojitě kladné funkce na  $\langle t_0, t_f \rangle$ . Potom řešení (1.2.1) existuje a je jednoznačné na celém intervalu  $\langle t_0, t_f \rangle$ .*

Tento výsledek lze zobecnit pro rovnice s více zpožděními. Zvláště pro lineární ZDR,

$$\begin{cases} y'(t) = \sum_{i=0}^r A_i(t)y(t - \tau_i) + g(t), & t \geq t_0 \\ y(t) = \phi(t), & t \leq t_0, \end{cases} \quad (1.2.3)$$

kde  $A_i(t), i = 0, \dots, r$ ,  $g(t)$  a  $\phi(t)$  jsou libovolné spojitě funkce, platí globální existence, jednoznačnost a spojitá závislost na počátečních údajích.

Můžeme také použít teorii ODR pro určení existence a jednoznačnosti. Na intervalu  $[t_0, t_0 + \tau]$  počáteční problém ZDR (0.0.1) přechází na počáteční problém pro ODR:

$$\begin{cases} y'(t) = f(t, y(t), \phi(t - \tau)), & t \geq t_0, \\ y(t_0) = \phi(t_0) \end{cases}$$

a tedy spojitost funkce  $f(t, u, v)$  vzhledem k  $t$  a Lipschitzovská spojitost vzhledem k  $u$  a  $v$ , dohromady s Lipschitzovskou spojitostí  $\phi$  a  $\phi'$ , nám garantuje lokální existenci a jednoznačnost na intervalu  $[t_0, t_0 + \delta]$  pro nějaké  $\delta > 0$ .

Informaci o existenci řešení na konečném intervalu  $[t_0, t_f]$  můžeme získat pomocí po sobě jdoucí integrací na intervalech  $[t_0 + i\tau, t_0 + (i + 1)\tau]$ ,  $i = 0, 1, \dots$ , kde řešíme stále obyčejné diferenciální rovnice. Tato metoda se nazývá *metodou kroků* a je to jedna ze základních metod pro teoretickou analýzu a numerickou integraci ZDR nebo soustavy ZDR.

## 2 Obyčejné diferenciální rovnice - spojité numerické metody

Nechť  $\|\cdot\|$  je daná norma na  $\mathbb{R}^d$ . Uvažujme počáteční problém pro soustavu obyčejných diferenciálních rovnic (ODR)

$$\begin{cases} y'(t) = g(t, y(t)), & t_0 \leq t \leq t_f, \\ y(t_0) = y_0, \end{cases} \quad (2.0.1)$$

kde  $g(t, y) \in C^0([t_0, t_f] \times \mathbb{R}^d, \mathbb{R}^d)$  je spojitá funkce vzhledem k  $t$  a globálně Lipschitzovsky spojitá vzhledem k  $y$  v dané normě  $\|\cdot\|$ , tedy

$$\|g(t, y_1) - g(t, y_2)\| \leq L\|y_1 - y_2\| \quad \forall t \in [t_0, t_f] \text{ a } \forall y_1, y_2 \in \mathbb{R}^d, \quad (2.0.2)$$

pro nějakou Lipschitzovu konstantu  $L > 0$ . Dále mějme síť  $\Delta = \{t_0, \dots, t_N = t_f\}$  a odpovídající krok  $h_{n+1} = t_{n+1} - t_n$ ,  $n = 0, \dots, N - 1$ . Uvažujme obecnou jedнокrokovou metodu na této síti

$$y_{n+1} = \alpha_n y_n + h_{n+1} \Phi(y_n, g, \Delta_n), \quad (2.0.3)$$

pro  $n \geq 0$ , kde  $\Delta_n = \{t_n, t_{n+1}\}$  a funkce přírůstku  $\Phi$  splňuje globální Lipschitzovu podmínku vzhledem k  $y$  s Lipschitzovou konstantou  $\tilde{L}$  závisující pouze na funkci  $g(t, y)$ , tedy přesněji na konstantě  $L$  v (2.0.2). Předpokládáme, že koeficienty  $\alpha_n$  jsou stejnoměrně omezené vzhledem k  $n$ . Dále budeme předpokládat, že funkce přírůstku  $\Phi$  je spojitá vzhledem k funkci  $g$  v následujícím smyslu:

Existuje krok  $h_g > 0$  a konstanta  $\gamma_g > 0$  spojitě závislá na  $L$ , ale nezávislá na hodnotě  $y_n$  v bodech sítě a na bodech sítě  $\Delta_n$ , takže platí pro všechny kroky  $h_{n+1} \leq h_g$ :

$$\|\Phi(y_n, \tilde{g}, \Delta_n) - \Phi(y_n, g, \Delta_n)\| \leq \gamma_g \sup_{t_n \leq t \leq t_{n+1}} \|\tilde{g}(t, y) - g(t, y)\|, \quad (2.0.4)$$

pro všechna  $\tilde{g} \in C^0([t_0, t_f] \times \mathbb{R}^d, \mathbb{R}^d)$ . Všimněme si, že podmínka (2.0.4) je silnější než podmínka spojitosti vzhledem k funkci  $g$ , implikuje totiž pro fixní  $g$  sublineární růst funkce  $\Phi$  vzhledem k přírůstku  $\tilde{g} - g$ .

**Definice 2.0.1** *Nechť  $\eta(t)$  je po částech polynomiální funkce, jež je na každém intervalu  $[t_n, t_{n+1}]$  dána restrikcí interpolační funkce konstruované na obecně širším intervalu  $[t_{n-i_n}, t_{n+j_n+1}]$ ,  $i_n, j_n \geq 0$ , a sice tvaru*

$$\begin{aligned} \eta(t_n + \theta h_{n+1}) &= \beta_{n,1}(\theta) y_{n+j_n} + \dots + \beta_{n,j_n+i_n+1}(\theta) y_{n-i_n} \\ &\quad + h_{n+1} \Psi(y_{n+j_n}, \dots, y_{n-i_n}, \theta, g, \Delta'_n), \quad 0 \leq \theta \leq 1, \end{aligned} \quad (2.0.5)$$

kde  $\Delta'_n = \{t_{n-i_n}, \dots, t_{n+j_n}, t_{n+j_n+1}\}$ . Nechť navíc  $\eta(t)$  splňuje podmínku spojitosti:

$$\eta(t_n) = y_n \quad \text{a} \quad \eta(t_{n+1}) = y_{n+1}. \quad (2.0.6)$$

Pak  $\eta(t)$  nazveme spojitém prodloužením (interpolantem) ODR metody (2.0.3).

Pokud  $i_n = j_n = 0$ , pak interpolační postup je založen pouze na hodnotách patřících do jediného intervalu  $[t_n, t_{n+1}]$  a budeme  $\eta(t)$  označovat jako jedнокrokovou interpolaci. V opačném případě jako vícekrokovou interpolaci.

Každou ODR metodu (2.0.3) vybavenou jejím spojitém prodloužením budeme označovat jako spojitá ODR metoda.

Všimněme si, že pro  $j_n > 0$  interpolace (2.0.5) nemůže být spočítána zároveň s  $y_{n+1}$  pomocí diskrétní formule (2.0.5), ale pouze až diskrétní integrace postoupí k nadcházejícímu uzlovému bodu  $t_{n+j_n+1}$ .

**Poznámka 2.0.2** Protože podmínka spojitosti (2.0.6) musí platit pro každou rovnici, dostáváme tedy hlubší vztah mezi ODR metodou (2.0.3) a jejím spojitým prodloužením (2.0.5). Tedy spojitě prodloužení může být vyjádřeno pomocí (2.0.5) kde pro proměnné platí

$$\beta_{n,j}(0) = \begin{cases} 1 & \text{pro } j = 1 + j_n, \\ 0 & \text{jinak,} \end{cases}$$

$$\Psi(y_{n+j_n}, \dots, y_{n-i_n}, 0, g, \Delta'_n) = 0$$

a

$$\beta_{n,j}(1) = \begin{cases} \alpha_{n,j-j_n} & \text{pro } 1 + j_n \leq j \leq 1 + j_n, \\ 0 & \text{jinak,} \end{cases}$$

$$\Psi(y_{n+j_n}, \dots, y_{n-i_n}, 1, g, \Delta'_n) = \Phi(y_n, g, \Delta_n).$$

Speciálně pro  $\theta = 1$  formule (2.0.5) přejde na (2.0.3).

Předpokládáme, že funkce přírůstku  $\Psi$  splňuje globální Lipschitzovu podmínku vzhledem k  $y$  s Lipschitzovou konstantou závislou pouze na konstantě  $L$  v (2.0.2). Polynomické koeficienty  $\beta_{n,i}(\theta)$  jsou stejnoměrně omezené vzhledem k  $\theta$  a  $\Delta'_n$ . Podobně jako u diskrétních metod předpokládáme, že funkce přírůstku  $\Psi$  je spojitá vzhledem k funkci  $g$  v následujícím smyslu.

Existuje krok  $h_g > 0$  a konstanta  $\delta_g > 0$ , které jsou spojitě závislé na  $L$ , ale nezávislé na uzlových bodech  $y_{n-i_n}, \dots, y_{n+j_n}$ , nezávislé na bodech sítě  $\Delta'_n$  a na  $\theta \in [0, 1]$ , tak že platí

$$\begin{aligned} & \|\Psi(y_{n+j_n}, \dots, y_{n-i_n}, \theta, \tilde{g}, \Delta'_n) - \Psi(y_{n+j_n}, \dots, y_{n-i_n}, \theta, g, \Delta'_n)\| \\ & \leq \delta_g \sup_{t_{n-i_n} \leq t \leq t_{n+j_n+1}} \|\tilde{g}(t, y) - g(t, y)\|, \end{aligned} \quad (2.0.7)$$

pro všechny kroky  $h_{n-i_n+1}, \dots, h_{n+j_n+1} \leq h_g$  a pro všechna  $\tilde{g} \in C^0([t_0, t_f] \times \mathbb{R}^d, \mathbb{R}^d)$ .

Pro aplikaci numerických metod budeme předpokládat následující vlastnost pro sítě  $\Delta$ . Existuje kladná konstanta  $\Omega$ , nezávislá na síti  $\Delta$ , taková že platí:

$$\Omega^{-1}h_{n+1+i} \leq h_{n+1} \leq \Omega h_{n+1+i}, \quad i = -i_n, +1, \dots, j_n$$

pro každé  $n$ .

Všimněme si, že pro jednokrokové interpolanty ( $i_n = j_n = 0$ ) neplatí pro síť žádné omezení.

V některých případech interpolace přirozeně vychází ze samotné metody. Toto je případ *Adamsových* a *kolokačních* metod. I přesto je interpolace nezávislá na diskrétní metodě.

**Definice 2.0.3** Řekněme, že ODR metoda (2.0.3) má řád (diskrétní řád)  $p$ , pokud  $p \geq 1$  je největší přirozené číslo takové, že pro všechny spojitě funkce  $g$  v (2.0.1) třídy  $C^p$  a pro všechny body sítě platí

$$\|z_{n+1}(t_{n+1}) - \tilde{y}_{n+1}\| = O(h_{n+1}^{p+1})$$

stejněměrně vzhledem k  $y_n^*$  v každé ohraničené podmnožině  $\mathbb{R}^d$  a vzhledem k  $n = 0, \dots, N-1$ , kde  $z_{n+1}(t)$  je lokální řešení lokálního problému

$$\begin{cases} z'_{n+1}(t) = g(t, z_{n+1}(t)), & t_n \leq t \leq t_{n+1}, \\ z_{n+1}(t_n) = y_n^* \end{cases} \quad (2.0.8)$$

a

$$\tilde{y}_{n+1} = \alpha_n z_{n+1}(t) + h_{n+1} \Phi(z_{n+1}(t_n), g, \Delta_n). \quad (2.0.9)$$

**Definice 2.0.4** Řekněme, že interpolant (2.0.5) ODR metody (2.0.3) má stejnoměrný řád  $q$  pokud  $q \geq 1$  je největší přirozené číslo takové, že pro všechny spojité funkce  $g$  z (2.0.1) třídy  $C^q$  a pro všechny síťové body platí

$$\max_{t_n \leq t \leq t_{n+1}} \|z_{n+1}(t) - \tilde{\eta}(t)\| = O(h_{n+1}^{q+1}), \quad (2.0.10)$$

kde

$$\begin{aligned} \tilde{\eta}(t_n + \theta h_{n+1}) &= \beta_{n,1}(\theta) z_{n+1}(t_{n+j_n}) + \dots + \beta_{n,j_n+i_n+1}(\theta) z_{n+1}(t_{n-i_n}) \\ &\quad + h_{n+1} \Psi(z_{n+1}(t_{n+j_n}), \dots, z_{n+1}(t_{n-i_n}), \theta, g, \Delta'_n). \end{aligned} \quad (2.0.11)$$

**Poznámka 2.0.5** Z důvodu podmínky spojitosti (2.0.6) je zřejmé, že stejnoměrný řád  $q$  interpolantu (2.0.5) a diskrétní řád  $p$  metody (2.0.3) splňují nerovnost

$$1 \leq q \leq p.$$

**Věta 2.0.6** Necht' jednokroková ODR metoda (2.0.3) má diskrétní řád  $p \geq 1$ .

Pokud

- metoda je  $D$ -stabilní (stabilní ve smyslu Dahlquistova), tedy pro každé  $n$  platí

$$|\alpha_n| \leq 1; \quad (2.0.12)$$

- spojitá funkce  $g(t, y)$  v (2.0.1) je třídy  $C^p$ ,

potom ODR metoda (2.0.3) má globální řád  $p$  na jakémkoliv ohraničeném intervalu  $[t_0, t_f]$ , tedy

$$\max_{1 \leq n \leq N} \|y(t_n) - y_n\| = O(h^p), \quad (2.0.13)$$

kde  $h = \max_{1 \leq n \leq N} h_n$ .

Dále pokud interpolant (2.0.5) má stejnoměrný řád  $q$ , potom ODR metoda (2.0.3), (2.0.5) má globální stejnoměrný řád  $q' = \min\{p, q + 1\}$ , tedy:

$$\max_{t_0 \leq t \leq t_f} \|y(t) - \eta(t)\| = O(h^{q'}). \quad (2.0.14)$$

Důkaz můžeme nalézt v [3] (3.kapitola, strana 44-46).

**Poznámka 2.0.7** Pokud je ODR metoda (2.0.3) implementována s konstantním krokem  $h$ , potom koeficienty  $\alpha_n$  jsou nezávislé na síti  $\Delta$ , a proto všechny koeficienty  $\alpha_n = \alpha$ . Tedy  $|\alpha| \leq 1$  zajistí  $D$ -stabilitu metody.

### 3 Zpožděné diferenciální rovnice - numerické řešení

#### 3.1 Feldsteinova metoda

Mějme počáteční problém pro diferenciální rovnici s konstantním zpožděním

$$\begin{cases} y'(t) = (f(t, y(t), y(t - \tau))), & t_0 \leq t \leq t_f, \\ y = \phi(t), & t \leq t_0. \end{cases} \quad (3.1.1)$$

Vytvoříme množinu  $\Delta = \{t_0, t_1, \dots, t_n, \dots, t_N = t_f\}$  síťových bodů (tzv. síť) takovou, aby platilo pro  $\forall t_n \in \Delta$  buď  $t_n - \tau < t_0$  nebo  $t_n - \tau \in \Delta$ . Potom každá diskretní metoda, která používá pouze síťové body může být implementována přímo pro zpožděné diferenciální rovnice. Jako příklad uvedeme *dopřednou Eulerovu* metodu pro rovnici (3.1.1)

$$y_{n+1} = y_n + h_{n+1}f(t_n, y_n, y_q),$$

pro nějaké celé číslo  $q < n$  kde  $y_q = y(t_n - \tau)$ .

Pokud nechceme sestavit síť v závislosti na zpoždění, lze použít vyčíslení zpožděného členu  $y(t - \tau)$  pomocí aproximace. Uvedeme tedy Feldsteinovu formuli pro řešení rovnice (3.1.1).

Zaveďme krok  $h = \frac{t_f - t_0}{m}$ , kde  $m$  je přirozené číslo. Pro každý bod  $t_n = t_0 + nh$  definujeme celočíselnou funkci  $q(n) = \lfloor \frac{t_n - \tau - t_0}{h} \rfloor$  a funkci  $r(n) = \frac{t_n - \tau - t_0}{h} - q(n)$ . Všimněme si, že platí

- $t_n - \tau = t_{q(n)} + hr(n)$ ;
- $0 \leq r(n) < 1$ ;
- $0 \leq q(n) \leq \frac{t_n - \tau - t_0}{h} < n$ .

První Feldsteinova formule je tvaru

$$\begin{cases} y_{n+1} = y_n + hf(t_n, y_n, z_n), \\ y_0 = y(t_0), \end{cases}$$

kde

$$z_n = y_{q(n)} + hr(n)f(t_{q(n)}, y_{q(n)}, z_{q(n)}) = (1 - r(n))y_{q(n)} + r(n)y_{q(n)+1}.$$

To odpovídá po částech lineární interpolaci  $y(t_n - \tau)$ . Pokud pro  $n$  platí  $t_n - \tau < t_0$ , tak potom  $z_n = \phi(t_n - \tau)$ .

Místo po částech lineární interpolace Feldstein navrhl druhou formuli tvaru

$$z_n = y_{q(n)},$$

která odpovídá po částech konstantní aproximaci  $y(t_n - \tau)$ , kde  $t_n - \tau \geq t_0$ .

Feldsteinova metoda je implementovaná s první i druhou formulí v Matlabovském skriptu "Feldstein.mlx"(viz. Příloha).

## 3.2 Bellmanova metoda kroků

Jedná se o elegantní přístup k řešení numerického řešení počátečního problému (3.1.1) bez nutnosti interpolace.

Je zřejmé, že pro konstantní zpoždění body nespojitosti jsou  $\xi_k = t_0 + k\tau$ . V prvním makro intervalu  $[t_0, t_0 + \tau]$  má ZDR (3.1.1) tvar

$$\begin{cases} y'(t) = f(t, y(t), \phi(t - \tau)), \\ y(t_0) = \phi(t_0). \end{cases}$$

V druhém makro intervalu  $[t_0 + \tau, t_0 + 2\tau]$  můžeme definovat  $y_1(t) = y(t - \tau)$  a  $y_2(t) = y(t)$ , tedy můžeme napsat ZDR (3.1.1) jako  $2d$ -dimenzionální systém ODR rovnic

$$\begin{cases} y_1'(t) = f(t - \tau, y_1(t), \phi(t - 2\tau)), \\ y_2'(t) = f(t, y_2(t), y_1(t)), \\ y_1(t_0 + \tau) = \phi(t_0), \\ y_2(t_0 + \tau) = y(t_0 + \tau). \end{cases}$$

Obecně v makro intervalu  $[t_0 + (k - 1)\tau, t_0 + k\tau]$  můžeme napsat ZDR (3.1.1) jako  $kd$ -dimenzionální systém ODR rovnic

$$\begin{cases} y_i'(t) = f(t - (k - i)\tau, y_1(t), y_{i-1}(t)), & i = 1, \dots, k, \\ y_i(t_0 + (k - 1)\tau) = y(t_0 + (i - 1)\tau), & i = 1, \dots, k, \end{cases} \quad (3.2.1)$$

kde položíme  $y_0(t) = \phi(t - k\tau)$  a  $y_i(t) = y(t - (k - i)\tau)$ ,  $i = 1, \dots, k$ .

Přechod z  $k$  na  $k + 1$  v (3.2.1) znamená posunutí integračního intervalu z  $[t_0 + (k - 1)\tau, t_0 + k\tau]$  na  $[t_0 + k\tau, t_0 + (k + 1)\tau]$  a prodloužení řešení z  $[t_0, t_0 + k\tau]$  do  $[t_0, t_0 + (k + 1)\tau]$  pomocí přidání členu  $y_{k+1}(t) = y(t)$ . Proto tedy můžeme použít numerickou metodu pro ODRn a řešit pro zvětšující se  $k$  větší a větší systém (3.2.1). Pro každý krok  $k$ , numerické řešení  $kd$ -dimenzionálního systému (3.2.1) cílí na poskytnutí aproximace hodnoty  $y_k(t_0 + k\tau) = y(t_0 + k\tau)$ , která bude použita jako počáteční hodnota pro nový člen  $y_{k+1}(t) = y(t)$  v dalším kroku. Proces skončí v momentě, kdy pro nějaké  $k$  platí  $t_0 + k\tau \geq t_f$ .

Tímto způsobem musíme řešit systém rostoucí dimenze a v principu vypočítáváme mnohokrát stejné části řešení (v předchozích intervalech). Na druhou stranu redukcí ZDR (3.1.1) na systém ODR rovnic se vyhneme typickým komplikacím souvisejícím s přítomností zpožděného argumentu  $t - \tau$ , tedy ukládání a interpolování vypočteného řešení na intervalu  $[t_0, t_f - \tau]$ .

Všimněme si, že v každém kroku jakákoliv numerická metoda řádu  $p$  poskytne numerickou aproximaci řádu  $p$  pro jakoukoliv poruchu počátečních podmínek v (3.2.1) řádu většího či rovno  $p$ . Proto tedy v dalším kroku musíme vyřešit počáteční problém s poruchou počátečních podmínek stejného řádu jako metoda. Celý proces tedy poskytne řešení řádu  $p$  pro konečný počet kroků.

Bellmanova metoda kroků je implementovaná s dopřednou Eulerovou metodou v Matlabovském skriptu "Bellman\_forward\_euler.mlx" (viz. Příloha).

## 3.3 Standardní přístup pomocí spojitých ODR metod

Mějme počáteční problém pro ZDR s konstantním zpožděním:

$$\begin{cases} y'(t) = (f(t, y(t), y(t - \tau))), & t \geq t_0, \\ \phi(t) = y_0, & t \leq t_0, \end{cases} \quad (3.3.1)$$

kde  $f(t, u, v) \in C^0([t_0, t_f] \times \mathbb{R}^d, \mathbb{R}^d)$  je spojitá funkce vzhledem k  $t$  a globálně Lipschitzov-sky spojitá vzhledem k  $u$  a  $v$  v dané normě  $\|\cdot\|$  v  $\mathbb{R}^d$ , tedy:

$$\|f(t, u_1, v_1) - f(t, u_2, v_2)\| \leq L\|u_1 - u_2\| + M\|v_1 - v_2\|,$$

pro všechna  $t \in [t_0, t_f]$ ,  $u_1, u_2, v_1, v_2 \in \mathbb{R}^d$  a pro nějaké Lipschitzovy konstanty  $L > 0$  a  $M > 0$ .

Nechť síť  $\Delta = \{t_0, t_1, \dots, t_n, \dots, t_N = t_f\}$  obsahuje všechny body nespojitosti řádu  $\leq p$  v  $[t_0, t_f]$  (viz. Definice 1.1.2). Označíme je  $\xi_0 = t_0 < \xi_1 < \dots < \xi_s < \xi_{s+1} = t_f$ . Na každém intervalu  $[\xi_0, \xi_1], [\xi_1, \xi_2], \dots, [\xi_s, t_f]$  je řešení  $y(t)$  třídy alespoň  $C^{p+1}$ . Potom posloupnost počátečních problémů

$$\begin{cases} z'(t) = f(t, z(t), \eta(t - \tau)), & \xi_i \leq t \leq \xi_{i+1}, \\ z(\xi_i) = \eta(\xi_i), \end{cases}$$

kde  $i = 0, \dots, s$ , je vyřešena pomocí metody (2.0.3) a následně interpolována pomocí (2.0.5), kde  $\eta(s)$  je dána počáteční funkcí  $\phi(s)$  pro  $s \leq t_0$  nebo pomocí již vypočteného numerického řešení pro  $s > t_0$ .

**Algoritmus 3.3.1 (Metoda kroků)** 1. Najděme všechny body nespojitosti řádu  $\leq p$  (pro konstantní zpoždění je  $\xi_i = t_0 + i\tau$ )

$$\xi_1, \dots, \xi_s < t_f$$

a položme  $\xi_0 = t_0, \xi_{s+1} = t_f$ ;

2. Vyřešme rovnici:

$$\begin{cases} z'(t) = f(t, z(t), \phi(t - \tau)) & \xi_0 \leq t \leq \xi_1, \\ z(\xi_0) = \phi(\xi_0) \end{cases}$$

*jakoukoliv diskrétní numerickou ODR metodou;*

3. Pro  $i = 1, \dots, s$ :

- spočítejme a uložíme spojitě prodloužení  $\eta(t)$  pro  $t \in [\xi_{i-1}, \xi_i]$  pomocí interpolace hodnot ze stejného makro intervalu  $[\xi_{i-1}, \xi_i]$
- vyřešme počáteční úlohu:

$$\begin{cases} z'(t) = f(t, z(t), \eta(t - \tau)), & \xi_i \leq t \leq \xi_{i+1}, \\ z(\xi_i) = \eta(\xi_i) \end{cases}$$

*jakoukoliv diskrétní numerickou ODR metodou;*

4. Vykresleme aproximaci řešení  $\eta(t)$  na intervalu  $[\xi_0, \xi_{s+1}]$ ;

5. Konec;

Algoritmus je implementovaný pro implicitní Eulerovu metodu a lichoběžníkovou metodu v Matlabovském skriptu "metoda\_kroku\_implicitni\_metody.mlx"(viz. Příloha).



Nyní uvedeme tvrzení o řádu přesnosti numerické metody. Důkaz lze najít v [3] (4.kapitola, strana 64-69).

**Věta 3.3.2** *Uvažujme zpožděnou diferenciální rovnici (3.3.1) kde  $f(t, y, x)$  je funkce třídy  $C^p$  v  $[t_0, t_f] \times \mathbb{R}^d \times \mathbb{R}^d$  a počáteční funkce  $\phi(t)$  je třídy  $C^p$ . Dále předpokládejme:*

- *Sít  $\Delta = \{t_0, t_1, \dots, t_n, \dots, t_N = t_f\}$  obsahuje všechny body nespojitosti řádu  $\leq p$  (viz. (1.1.2) ležící v intervalu  $[t_0, t_f]$ , označené jako  $\xi_1, \dots, \xi_s < t_f$ ;*
- *ODR metoda (2.0.3) má řád  $p$ , splňuje podmínku  $D$ -stability (2.0.12) a pro  $k > 1$  je restartována po dosažení každého bodu nespojitosti  $\xi_i$ ,  $i = 0, 1, \dots, s$ , pomocí metody řádu  $\geq p - 1$ ;*
- *Interpolant (2.0.5) má stejnoměrný řád  $q$ ;*
- *Pro každé  $n$ , interval  $[t_{n-i_n}, t_{n+j_n+i}]$ , kde se odehrává interpolace, je součástí  $[\xi_i, \xi_{i+1}]$  pro nějaký index  $0 \leq i \leq s$ ;*

*Potom má výsledná metoda kroků diskrétní globální řád a stejnoměrný globální řád  $q' = \min\{p, q + 1\}$ , tedy:*

$$\max_{1 \leq n \leq N} \|y(t_n) - y_n\| = O(h^{q'})$$

*a*

$$\max_{t_0 \leq t \leq t_f} \|y(t) - \eta(t)\| = O(h^{q'}),$$

*kde  $h = \max_{1 \leq n \leq N} h_n$ .*

## 4 Spojité Rungeovy-Kuttovy metody pro obyčejné diferenciální rovnice

Mějme počáteční problém pro ODR:

$$\begin{cases} y'(t) = g(t, y(t)), & t_0 \leq t \leq t_f, \\ y(t_0) = y_0. \end{cases} \quad (4.0.1)$$

Mějme danou síť  $\Delta = \{t_0, t_1, \dots, t_n, \dots, t_N = t_f\}$ ,  $\nu$ -stupňová Rungeova-Kuttova (RK) metoda pro řešení počátečního problému (4.0.1) má tvar (v tzv. *Y notaci*):

$$Y_{n+1}^i = y_n + h_{n+1} \sum_{j=1}^{\nu} a_{ij} g(t_{n+1}^j, Y_{n+1}^j), \quad i = 1, \dots, \nu, \quad (4.0.2)$$

$$y_{n+1} = y_n + h_{n+1} \sum_{i=1}^{\nu} b_i g(t_{n+1}^i, Y_{n+1}^i), \quad (4.0.3)$$

kde  $t_{n+1}^i = t_n + c_i h_{n+1}$ ,  $c_i = \sum_{j=1}^{\nu} a_{ij}$ ,  $i = 1, \dots, \nu$ ,  $h_{n+1} = t_{n+1} - t_n$ , kde  $\nu$  je počet *stupňů*. Koeficienty  $b_i$  se nazývají *váhy* kvadraturní formule (4.0.3) a koeficienty  $c_i$  označujeme jako *abscisy*, pro většinu obvyklých metod platí  $c_i \in [0, 1]$ . Protože je RK metoda (4.0.2), (4.0.3) charakterizována pomocí vah  $b_i$  a matice koeficientů  $A = [a_{ij}]_{i,j=1}^{\nu}$ , budeme ji označovat jako  $(A, b)$  RK metodu. Koeficienty RK metody budeme zapisovat do tzv. *Butcherovy tabulky*

$$\begin{array}{c|c} C & A \\ \hline & B \end{array}$$

kde  $C = [c_1, c_2, \dots, c_{\nu}]^T$  a  $B = [b_1, b_2, \dots, b_{\nu}]$ .

Můžeme také použít místo (4.0.2), (4.0.3) ekvivalentní formu (v tzv. *K notaci*)

$$K_{n+1}^i = g\left(t_{n+1}^i, y_n + h_{n+1} \sum_{j=1}^{\nu} a_{ij} K_{n+1}^j\right), \quad i = 1, \dots, \nu,$$

$$y_{n+1} = y_n + h_{n+1} \sum_{i=1}^{\nu} b_i K_{n+1}^i.$$

Všimněme si, že *K notaci* získáme z formule (4.0.2), (4.0.3) pomocí

$$K_{n+1}^i = g(t_{n+1}^i, Y_{n+1}^i), \quad i = 1, \dots, \nu.$$

Výpočetní složitost metody je hlavně určena počtem stupňů metody a koeficienty matice  $A$ . Je známo, že pokud je matice  $A$  dolní trojúhelníková s nulovými prvky na diagonále, potom se metoda nazývá *explicitní* a výpočetní náročnost je menší, zatímco v opačném případě (alespoň jeden prvek na diagonále či v horním trojúhelníku je nenulový) hovoříme o *implicitní* metodě a výpočetní náročnost je větší.

## 4.1 Spojité prodloužení RK metod

### 4.1.1 Interpolant první třídy

Interpolant, který je zkonstruován bez použití dalších stupňů, budeme označovat jako *interpolant první třídy* a výsledné spojité prodloužení  $\eta(t)$  je definováno na každém podintervalu sítě  $\Delta$  pomocí jednokrokové spojité kvadraturní formule

$$\eta(t_n + \theta h_{n+1}) = y_n + h_{n+1} \sum_{i=1}^{\nu} b_i(\theta) g(t_{n+1}^i, Y_{n+1}^i), \quad 0 \leq \theta \leq 1 \quad (4.1.1)$$

nebo v *K notaci*

$$\eta(t_n + \theta h_{n+1}) = y_n + h_{n+1} \sum_{i=1}^{\nu} b_i(\theta) K_{n+1}^i, \quad 0 \leq \theta \leq 1,$$

kde  $b_i(\theta)$  jsou polynomy vhodného stupně  $\leq \delta$  splňující:

$$b_i(0) = 0 \text{ a } b_i(1) = b_i, \quad i = 1, \dots, \nu \quad (4.1.2)$$

a tedy splňuje podmínku spojitosti (2.0.6).

### 4.1.2 Interpolant druhé třídy

Interpolant zkonstruovaný pomocí dalších stupňů nazýváme *interpolant druhé třídy* a spojité prodloužení je dáno

$$\eta(t_n + \theta h_{n+1}) = y_n + h_{n+1} \sum_{i=1}^s b_i(\theta) g(t_{n+1}^i, Y_{n+1}^i), \quad 0 \leq \theta \leq 1 \quad (4.1.3)$$

nebo v *K notaci*

$$\eta(t_n + \theta h_{n+1}) = y_n + h_{n+1} \sum_{i=1}^s b_i(\theta) K_{n+1}^i, \quad 0 \leq \theta \leq 1,$$

kde  $b_i(\theta)$  jsou opět polynomy vhodného stupně  $\leq \delta$  splňující podmínku spojitosti

$$b_i(0) = 0, \quad i = 1, \dots, s,$$

$$b_i(1) = b_i, \quad i = 1, \dots, \nu; \quad b_i(1) = 0, \quad i = \nu + 1, \dots, s. \quad (4.1.4)$$

Další  $s - \nu$  stupně jsou dány pomocí

$$Y_{n+1}^i = y_n + h_{n+1} \sum_{j=1}^s a_{ij} g(t_{n+1}^j, Y_{n+1}^j), \quad i = \nu + 1, \dots, s \quad (4.1.5)$$

nebo v *K notaci*

$$K_{n+1}^i = g\left(t_{n+1}^i, y_n + h_{n+1} \sum_{j=1}^s a_{ij} K_{n+1}^j\right), \quad i = \nu + 1, \dots, s. \quad (4.1.6)$$

Můžeme tedy zadefinovat rozšířenou matici koeficientů  $A'$ :

$$A' = \begin{bmatrix} A & 0 \\ [a_{ij}]_{i=\nu+1, j=1}^{s, \nu} & [a_{ij}]_{i, j=\nu+1}^s \end{bmatrix} \quad (4.1.7)$$

kde  $A = [a_{ij}]_{i, j=1}^\nu$ .

Celkové spojité Rungeovy-Kuttovy (SRK) metody (4.0.2), (4.0.3), (4.1.1) (s koeficienty  $(A, b(\theta))$ ) a (4.0.2), (4.0.3), (4.1.5), (4.1.3) (s koeficienty  $(A', b(\theta))$ ) jsou spojitém prodloužením RK metody (4.0.2), (4.0.3) (s koeficienty  $(A, b)$ ) a  $\delta$  bude značit *stupeň interpolantu*. Oproti tomu RK metoda (4.0.2), (4.0.3) (s koeficienty  $(A, b)$ ) bude nazývána jako diskrétní RK metoda.

Je vhodné poznamenat, že obecně neplatí  $\eta(t_n + c_i h_{n+1}) \neq Y_{n+1}^i$ . Nicméně rovnost platí pro každou pravou stranu  $g(t, y)$ , pokud  $b_j(c_j) = a_{ij}$ , toto vychází z porovnání (4.1.3) a (4.1.5). Máme tedy

$$\{\eta(t_n + c_i h_{n+1}) = Y_{n+1}^i \quad \forall i\} \iff \{b_i(c_j) = a_{ji} \quad \forall i, j\}. \quad (4.1.8)$$

Interpolant první nebo druhé třídy tedy určuje matici  $\tilde{B}$  a její prvky jsou  $b_{ij} = b_j(c_i)$ .

**Definice 4.1.1** *SRK metoda se nazývá přirozená, pokud platí  $A = \tilde{B}$  ( $A' = \tilde{B}$ ).*

Funkce přírůstku  $\Phi$  a  $\Psi$  z (2.0.3) a (2.0.5) pro SRK metody jsou tvaru

$$\Phi(y_n, g, t_n, t_{n+1}) = \sum_{i=1}^{\nu} b_i g(t_{n+1}^i, Y_{n+1}^i)$$

a

$$\Psi(y_n, \theta, g, t_n, t_{n+1}) = \sum_{i=1}^s b_i(\theta) g(t_{n+1}^i, Y_{n+1}^i).$$

**Tvrzení 4.1.2** *Funkce přírůstku  $\Phi$  RK metody (4.0.2), (4.0.3) a  $\Psi$  interpolantu (4.1.1) nebo (4.1.3), splňují podmínku spojitosti (2.0.4) a (2.0.7).*

Důkaz můžeme nalézt v [3] (5.kapitola, strana 113-114).

**Věta 4.1.3** *Pokud pro SRK metodu platí Věta 2.0.6 a navíc platí, že interpolant je po částech polynomiální funkce řádu  $\delta \geq q$  a funkce  $g(t, y)$  v (4.0.1) je třídy  $C^{\max\{\delta, p\}}$ , potom následující odhady konvergence, ohraničenosti a neohraničenosti platí pro všechny derivace funkce globální chyby*

$$\max_{t_0 \leq t \leq t_f} \|z_{n+1}^{(j)}(t) - \eta^{(j)}(t)\| = O(h_{n+1}^{q+1-j}), \quad j = 0, \dots, \delta, \quad (4.1.9)$$

kde derivace  $\eta(t)$  v uzlových bodech jsou myšleny v pravém/levém smyslu.

Důkaz lze nalézt v [3] (5.kapitola, strana 115-116).

**Věta 4.1.4** *Předpokládejme, že RK metoda (4.0.2), (4.0.3) má spojité prodloužení  $\eta(t)$  řádu  $q$  a stupně  $d > q$ . Potom existuje jiné spojité prodloužení  $\tilde{\eta}(t)$  řádu  $q$  a stupně také  $q$ .*

Důkaz lze nalézt v [3] (5.kapitola, strana 116-117).

**Poznámka 4.1.5** z Vět 4.1.3 a 4.1.4 je patrné, že požití interpolantu stupně vyššího než  $q$  je zbytečné. Interpolanty stupně  $\delta > q+1$  jsou dokonce nebezpečné, protože derivace řádu  $k$ ,  $q+2 \leq k \leq \delta$  mohou divergovat pro  $h \rightarrow 0$ . Kvůli těmto důvodům budeme předpokládat, že spojitě prodloužení řádu  $q$  bude vždy tvořeno interpolanty stupně  $\delta = q$ .

Zabývejme se nyní otázkami ohledně interpolantů první (4.1.1) a druhé třídy (4.1.3).

**Otázka 1** Jaký je nejvyšší možný stejnoměrný řád RK metody řádu  $p$  za použití interpolantu první třídy?

**Otázka 2** Uvažujme RK metodu diskrétního řádu  $p$  a spojitěho prodloužení  $q < p$ . Kolik je potřeba dalších stupňů, abychom dosáhli stejnoměrného řádu  $p-1$  nebo dokonce  $p$ ?

**Věta 4.1.6** Předpokládejme, že RK metoda (4.0.2), (4.0.3) má spojitě prodloužení  $\eta(t)$  tvaru (4.1.3). Potom stejnoměrný řád  $q$  nemůže být vyšší než  $s^*$ , tedy počet různých abscis prodloužené RK metody reprezentované pomocí (4.1.7).

Důkaz lze nalézt v [3] (5.kapitola, strana 118).

## 4.2 Analýza interpolantu první třídy

Obecná analýza stejnoměrného řádu pro spojitě prodloužení (4.1.1) je založena na tom, že pro  $0 < \theta \leq 1$  se na prodloužení můžeme dívat jako na diskrétní metodu  $(\frac{A}{\theta}, \frac{b(\theta)}{\theta})$  s krokem  $\theta h_{n+1}$ . Okamžitě tedy dostaneme podmínky pro polynomy  $b_i(\theta)$  z dobře známých podmínek pro RK metody. Podmínky pro řád až  $p=4$  jsou uvedeny v tabulce 1.

Za účelem odpovědět na **Otázku 1**, každá metoda musí být analyzovaná individuálně pomocí kontroly podmínek řádu. Obecně můžeme dát pouze částečnou odpověď pomocí následující věty:

**Věta 4.2.1** Každá RK metoda (4.0.2), (4.0.3) řádu  $p \geq 1$  má spojitě prodloužení  $\eta(t)$  řádu (a stupně)  $q = 1, \dots, \lfloor \frac{p+1}{2} \rfloor$ .

Důkaz lze nalézt v [3] (5.kapitola, strana 118-120).

**Věta 4.2.2** Pokud má RK metoda (4.0.2), (4.0.3) spojitě prodloužení  $\eta(t)$  řádu (a stupně)  $q \geq 2$ , potom má také jiné spojitě prodloužení  $\tilde{\eta}(t)$  řádu (a stupně)  $\tilde{q}$ ,  $\tilde{q} \leq q-1$ .

Důkaz lze nalézt v [3] (5.kapitola, strana 120).

Závěrem tedy můžeme říct, že obecná odpověď na **Otázku 1** je taková, že pouze interpolanty do řádu  $\lfloor \frac{p+1}{2} \rfloor$  zaručeně existují. Na druhou stranu je možné (za použití interpolantu první třídy (4.1.1)), aby maximální stejnoměrný řád byl větší než  $\lfloor \frac{p+1}{2} \rfloor$  nebo dokonce  $= p$ .

**Definice 4.2.3** RK metodu (4.0.2), (4.0.3) diskrétního řádu  $p$  nazveme jako *superkonvergentní*, pokud maximální stejnoměrný řád  $q$  je dosažitelný pomocí interpolantů první třídy (4.1.1) stupně  $\leq p-1$ .

Jinými slovy, *superkonvergence* je dosažena v koncových bodech každého intervalu  $[t_n, t_{n+1}]$  vzhledem k maximální stejnoměrné přesnosti řádu  $q$ . Také může nastat situace, že interpolant dosáhne vyššího řádu  $p' > q$  v nějakých bodech uvnitř intervalu  $[t_n, t_{n+1}]$ . Tyto body budeme nazývat *superkonvergentní vnitřní body*.

Řád	Podmínky
1	$\sum_{i=1}^{\nu} b_i(\theta) = \theta$
2	$\sum_{i=1}^{\nu} b_i(\theta)c_i = \frac{1}{2}\theta^2$
3	$\sum_{i=1}^{\nu} b_i(\theta)c_i^2 = \frac{1}{3}\theta^3$ $\sum_{i,j=1}^{\nu} b_i(\theta)a_{ij}c_i = \frac{1}{6}\theta^3$
4	$\sum_{i=1}^{\nu} b_i(\theta)c_i^3 = \frac{1}{4}\theta^4$ $\sum_{i,j=1}^{\nu} b_i(\theta)c_i a_{ij}c_j = \frac{1}{8}\theta^4$ $\sum_{i,j=1}^{\nu} b_i(\theta)a_{ij}c_i^2 = \frac{1}{12}\theta^4$ $\sum_{i,j,k=1}^{\nu} b_i(\theta)a_{ij}a_{jk}c_i = \frac{1}{24}\theta^4$

Tabulka 1: Podmínky řádu pro spojité RK metody

#### 4.2.1 Kolokační metody

Jednokroková kolokační metoda je zavedena následovně: Vyberme  $\nu$  různých abscis  $c_1, \dots, c_\nu \in [0, 1]$  a v každém intervalu  $[t_n, t_{n+1}]$  vypočteme polynom  $\eta(t)$  stupně  $\delta \leq \nu$  splňující

$$\eta'(t_{n+1}^i) = g(t_{n+1}^i, \eta(t_{n+1}^i)), \quad i = 1, \dots, \nu, \quad \eta(t_n) = y_n. \quad (4.2.1)$$

Metoda může být přepsána jako spojitá implicitní RK metoda (4.0.2), (4.0.3), kde:

$$a_{ij} = \int_0^{c_i} \ell_j(\xi) d\xi, \quad i, j = 1, \dots, \nu,$$

$$b_i = \int_0^\theta \ell_i(\xi) d\xi, \quad i = 1, \dots, \nu,$$

kde  $\ell_i(\xi)$  jsou Lagrangeovy polynomické koeficienty  $\ell_i(\xi) = \prod_{k=1, k \neq i}^{\nu} \frac{\xi - c_k}{c_i - c_k}$ .

Všimněme si, že platí  $b_i(c_j) = a_{ji}$ , a proto každá kolokační metoda je přirozená SRK metoda (viz. 4.1.1).

Je známo, že pro jakoukoliv volbu abscis  $c_1, \dots, c_\nu \in [0, 1]$ , kolokační metoda má řád  $p \geq \nu$  a interpolant (4.1.1) má stejnoměrný řád  $q = \nu$ . Tedy podle Věty 2.0.6 kolokační metoda je spojitá RK metoda globálního stejnoměrného řádu  $q' = \nu$  (pokud  $p = \nu$ ) a nebo  $q' = \nu + 1$  (pokud  $p > \nu$ ). V tomto smyslu je kolokační metoda optimální, pokud chceme dosáhnout maximálního možného stejnoměrného řádu pro daný počet stupňů. Zejména v případě, kdy abscissy jsou posunuté kořeny Legendreova ortogonálního polynomu stupně  $\nu$ , potom má metoda řád  $p = 2\nu$ . To je nejznámější případ superkonvergence.

Nyní uvedeme nějaké příklady superkonvergentních kolokačních metod.

- Gaussovy metody (diskrétní řád  $p = 2\nu$ , stejnoměrný řád  $q = \nu$ , globální stejnoměrný řád  $q' = \min\{p, q + 1\}$ ):

$\nu = 1$  (*obdélníková metoda*):  $p = 2$ ,  $q = 1$ ,  $q' = 2$ ;

$$\begin{array}{c|c} \frac{1}{2} & \frac{1}{2} \\ \hline & 1 \end{array} \quad b_1(\theta) = \theta.$$

$\nu = 2$  (*Hammer-Hollingsworthova metoda*):  $p = 4$ ,  $q = 2$ ,  $q' = 3$ ;

$$\begin{array}{c|cc} \frac{1}{2} - \frac{\sqrt{3}}{6} & \frac{1}{4} & \frac{1}{4} - \frac{\sqrt{3}}{6} \\ \frac{1}{2} + \frac{\sqrt{3}}{6} & \frac{1}{4} + \frac{\sqrt{3}}{6} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array} \quad \begin{aligned} b_1(\theta) &= -\frac{\sqrt{3}}{2}\theta \left( \theta - 1 - \frac{\sqrt{3}}{3} \right), \\ b_2(\theta) &= \frac{\sqrt{3}}{2}\theta \left( \theta - 1 + \frac{\sqrt{3}}{3} \right). \end{aligned}$$

- Radau IIA metody (diskrétní řád  $p = 2\nu - 1$ , stejnoměrný řád  $q = \nu$ , globální stejnoměrný řád  $q' = \min\{p, q + 1\}$ ):

$\nu = 1$  (*implicitní Eulerova metoda*):  $p = 1$ ,  $q = 1$ ,  $q' = 1$ ;

$$\begin{array}{c|c} 1 & 1 \\ \hline & 1 \end{array} \quad b_1(\theta) = \theta.$$

$\nu = 2$  (*Ehleova metoda*):  $p = 3$ ,  $q = 2$ ,  $q' = 3$ ;

$$\begin{array}{c|cc} \frac{1}{3} & \frac{5}{12} & -\frac{1}{12} \\ 1 & \frac{3}{4} & \frac{1}{4} \\ \hline & \frac{3}{4} & \frac{1}{4} \end{array} \quad \begin{aligned} b_1(\theta) &= -\frac{3}{4}\theta(\theta - 2), \\ b_2(\theta) &= \frac{3}{4}\theta(\theta - \frac{2}{3}). \end{aligned}$$

- Lobattovy IIIA metody (diskrétní řád  $p = 2\nu - 2$ , stejnoměrný řád  $q = \nu$ , globální stejnoměrný řád  $q' = \min\{p, q + 1\}$ ):

$\nu = 2$  (*Lichoběžníková metoda*):  $p = 2$ ,  $q = 2$ ,  $q' = 2$ ;

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array} \quad \begin{aligned} b_1(\theta) &= -\frac{1}{2}\theta(\theta - 2), \\ b_2(\theta) &= \frac{1}{2}\theta. \end{aligned}$$

$\nu = 3$  (*Ehleova metoda*):  $p = 4$ ,  $q = 3$ ,  $q' = 4$ ;

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{5}{24} & \frac{1}{3} & -\frac{1}{24} \\ 1 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ \hline & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{array} \quad \begin{aligned} b_1(\theta) &= 2\theta\left(\frac{1}{3}\theta^2 - \frac{3}{4}\theta + \frac{1}{2}\right), \\ b_2(\theta) &= -4\theta^2\left(\frac{1}{3}\theta - \frac{1}{2}\right), \\ b_3(\theta) &= 2\theta^2\left(\frac{1}{3}\theta - \frac{1}{4}\right). \end{aligned}$$

Implementovanou Radau IIA Ehle metodu pro řešení ZDR můžeme nalézt v Příloze jako matlabovský program "Radau\_IIA\_Ehle\_metoda\_RK\_pro\_ZDR\_.mlx"

**Definice 4.2.4 (Přirozené spojité prodloužení)** Řekneme, že interpolant  $\eta(t)$  první třídy (4.1.1) řádu a stupně  $q$  je přirozené spojité prodloužení (PSP) RK metody (4.0.2), (4.0.3) řádu  $p$ , pokud polynomy  $b_i(\theta), i = 1, \dots, \nu$  jsou takové, že  $\eta(t)$  splňuje navíc podmínku asymptotické ortogonalit

$$\left\| \int_{t_n}^{t_{n+1}} G(t)[z'_{n+1}(t) - \eta'(t)] dt \right\| = O(h_{n+1}^{p+1}), \quad (4.2.2)$$

pro jakoukoliv dostatečně hladkou maticovou funkci  $G$ . Podmínku (4.2.2) splňuje  $\eta(t)$  stejnoměrně vzhledem k  $n = 0, \dots, N - 1$ .  $z_{n+1}$  je řešení lokálního problému (2.0.8).

Můžeme si všimnout, že PSP splňuje nutně také podmínku

$$\left\| \int_{t_n}^{t_{n+1}} G(t)[z_{n+1}(t) - \eta(t)] dt \right\| = O(h_{n+1}^{p+1}), \quad (4.2.3)$$

pro každou dostatečně hladkou maticovou funkci  $G$ .

Navíc díky (4.2.2) a (4.2.3) snadno získáme:

$$\max_{0 \leq n \leq N-1} \left\| \int_{t_n}^{t_{n+1}} G(x)[y^{(j)}(x) - \eta^{(j)}(x)] dx \right\| = O(h^{p+1}), \quad j = 0, 1, \quad (4.2.4)$$

pro každou dostatečně hladkou maticovou funkci  $G$ , kde  $h = \max_{1 \leq n \leq N} h_n$ .

Vzhledem k Větám 2.0.6 a 4.1.3, podmínka (4.2.4) udává další vlastnost PSP ohledně globální konvergence. Je dobré poznamenat, že kolokační polynom jednokrokové kolokační metody (4.2.1), je PSP stupně  $q = \nu$  (viz. [36]).

Díky existenci PSP a jejího dosažitelného stejnoměrného řádu platí následující věta.

**Věta 4.2.5** Pokud interpolant  $\eta(t)$  první třídy (4.1.1) řádu a stupně  $q$  je PSP RK metody (4.0.2), (4.0.3) řádu  $p$ , potom  $q > \lfloor \frac{p+1}{2} \rfloor$ .

Důkaz lze nalézt v [3] (5.kapitola, strana 124).

PSP RK metody není jednoznačné, může existovat více přirozených spojitých prodloužení. Obecně lze pro získání PSP použít podmínky z Tabulky 1 a podmínku ortogonalit (4.2.2).

Nyní uvedeme vybrané příklady PSP pro některé explicitní RK (ERK) metody.

- Jednostupňová ERK metoda řádu  $p = 1$  (Eulerova metoda):

$$\star \quad q = \lfloor \frac{p+1}{2} \rfloor = \nu = p = 1$$

$$b_1(\theta) = \theta;$$

- Dvoustupňové ERK metody řádu  $p = 2$ :

$$\star \quad q = \lfloor \frac{p+1}{2} \rfloor = 1$$

$$b_i(\theta) = b_i \theta \quad i = 1, 2;$$

$$\star \quad q = \nu = p = 2$$

$$b_1(\theta) = (b_1 - 1)\theta^2 + \theta,$$

$$b_2(\theta) = b_2\theta^2;$$

- Třístupňové ERK metody řádu  $p = 3$  (kde  $c_2, c_3 \neq 0$ ):



$$\star q = \lfloor \frac{p+1}{2} \rfloor = 2$$

$$b_i(\theta) = w_i\theta^2 + (b_i + w_i)\theta, \quad i = 1, 2,$$

kde

$$w_1 = -\frac{1}{2c_3} - (c_3 - c_2)\lambda,$$

$$w_2 = c_3\lambda,$$

$$w_3 = \frac{1}{2c_3} - c_2\lambda,$$

kde  $\lambda \in \mathbb{R}$ . Tady máme ukázkou jednoparametrové skupiny PSP, vidíme tedy příklad nejednoznačnosti PSP minimálního řádu  $q$ ;

$$\star q = \nu = p = 3$$

neexistuje PSP řádu  $q = 3$ ;

- Čtyř-stupňové ERK metody řádu  $p = 4$ :

$$\star q = \lfloor \frac{p+1}{2} \rfloor = 2$$

$$b_i(\theta) = 3(2c_i - 1)b_i\theta^2 + 2(2 - 3c_i)b_i\theta, \quad i = 1, 2, 3;$$

$$\star q = 3$$

$$b_1(\theta) = 2(1 - 4b_1)\theta^3 + 3(3b_1 - 1)\theta^2,$$

$$b_i(\theta) = 4(3c_i - 2)b_i\theta^3 + 3(3 - 4c_i)b_i\theta^2, \quad i = 2, 3, 4;$$

$$\star q = \nu = p = 4$$

neexistuje PSP řádu  $q = 4$ ;

Zabývejme se nyní na chvíli spojitými metodami vyššího řádu. Jedno z prvních spojitých prodloužení vyššího řádu pro RK metody vytvořil Sarafyan [30]. Jeho technika spočívá v použití vhodné lineární kombinace hodnot jednotlivých stupňů za účelem získání aproximace  $y(t)$  vysokého řádu v nějakém vnitřním bodě. Potom se sestaví Hermitův-Birkhoffův interpolant v těchto bodech, v případě krajních bodů sítě může být potřeba vypočítat nějaké další hodnoty. Sarafyan [30] původně uvedl množinu 8-stupňových RK metod řádu 6 se spojitým prodloužením stejnoměrného řádu 4. Následně vyvinul několik skupin RK metod vyššího řádu  $p$ , s interpolací řádu  $p$  nebo  $p - 1$  (viz. [32] a [31]). Nyní uveďme jako příklad 10-stupňovou metodu řádu 6 se spojitým prodloužením řádu 5, tato metoda je použita v kódu DKL6G vytvořeného Corwinem, Sarafyanem a Thompsonem [4]:

$$y_{n+1} = y_n + h_{n+1} \sum_{i=1}^{10} b_i(1)K_{n+1}^i \quad (4.2.5)$$

a interpolantem:

$$\eta(t_n + \theta h_{n+1}) = y_n + h_{n+1} \sum_{i=1}^{10} b_i(\theta)K_{n+1}^i, \quad 0 \leq \theta \leq 1,$$

kde

$$\begin{aligned}
b_1(\theta) &= \theta - \frac{2427}{500}\theta^2 + \frac{3586}{375}\theta^3 - \frac{1659}{200}\theta^4 + \frac{66}{25}\theta^5 - \frac{184}{3375}b_{10}(\theta), \\
b_2(\theta) &= 0 \\
b_3(\theta) &= \frac{708}{125}\theta^2 - \frac{1977}{125}\theta^3 + \frac{789}{50}\theta^4 - \frac{27}{5}\theta^5 + \frac{188}{1125}b_{10}(\theta), \\
b_4(\theta) &= \frac{87}{50}\theta^2 - \frac{243}{50}\theta^4 + \frac{201}{40}\theta^4 - \frac{9}{5}\theta^5 - \frac{4}{225}b_{10}(\theta), \\
b_5(\theta) &= -\frac{88}{25}\theta^2 + \frac{1226}{75}\theta^3 - 21\theta^4 + \frac{42}{5}\theta^5 - \frac{344}{675}b_{10}(\theta), \\
b_6(\theta) &= -\frac{21}{100}\theta^2 + \frac{21}{25}\theta^3 - \frac{21}{40}\theta^4 - \frac{49}{225}b_{10}(\theta), \\
b_7(\theta) &= \frac{228}{125}\theta^2 - \frac{1197}{125}\theta^3 + \frac{741}{50}\theta^4 - \frac{171}{25}\theta^5 - \frac{532}{1125}b_{10}(\theta), \\
b_8(\theta) &= -\frac{161}{250}\theta^2 + \frac{1127}{750}\theta^3 - \frac{161}{200}\theta^4 - \frac{644}{3375}b_{10}(\theta), \\
b_9(\theta) &= 2\theta^3 - 5\theta^4 + 3\theta^5 + \frac{8}{27}b_{10}(\theta), \\
b_{10}(\theta) &= \text{libovolná funkce } \theta
\end{aligned}$$

a s Butcherovou tabulkou:

0	0									
$\frac{1}{9}$	$\frac{1}{9}$	0								
$\frac{1}{6}$	$\frac{1}{24}$	$\frac{3}{24}$	0							
$\frac{1}{3}$	$\frac{1}{6}$	$-\frac{3}{6}$	$\frac{4}{6}$	0						
$\frac{1}{2}$	$\frac{1}{8}$	0	0	$\frac{3}{8}$	0					
$\frac{2}{3}$	$-\frac{8}{21}$	$\frac{6}{21}$	$\frac{24}{21}$	$-\frac{24}{21}$	$\frac{16}{21}$	0				
$\frac{5}{6}$	$\frac{65}{456}$	$\frac{15}{456}$	$-\frac{80}{456}$	$\frac{330}{456}$	$-\frac{160}{456}$	$\frac{210}{456}$	0			
1	$\frac{35}{322}$	$-\frac{81}{322}$	$\frac{240}{322}$	$-\frac{360}{322}$	$\frac{680}{322}$	$-\frac{420}{322}$	$\frac{228}{322}$	0		
1	$\frac{161}{3000}$	0	$\frac{684}{3000}$	$\frac{315}{3000}$	$\frac{680}{3000}$	$\frac{315}{3000}$	$\frac{684}{3000}$	$\frac{161}{3000}$	0	
$\frac{2}{3}$	$\frac{7}{135}$	0	$\frac{32}{135}$	$\frac{12}{135}$	$\frac{32}{135}$	$\frac{7}{135}$	0	0	0	0

Můžeme si všimnout, že spojitě prodloužení nesplňuje podmínku spojitosti (4.1.4) pro  $\theta = 1$ . Implementovanou metodu pro ZDR s volbou  $b_{10}(\theta) = \frac{\theta^5}{4}$  lze nalézt v Příloze jako matlabovský soubor "DKLAG6.mlx". Dále poznamenejme, že tato metoda v [3] (5.kapitola, strana 134-135) není uvedena stejně jako v [4], liší se hlavně některými koeficienty před  $\theta^5$  u jednotlivých polynomů  $b_i(\theta)$  přes stejnou volbu  $b_{10}(\theta) = \frac{\theta^5}{4}$ . Navíc předpis

$$y_{n+1} = y_n + h_{n+1} \sum_{i=1}^{10} b_i K_{n+1}^i$$

neposkytoval správné výsledky z důvodu chyby v koeficientech  $b_i$  a proto je zde neuvádíme.

Podobná technika byla vyvinuta později Hornem [22], který získal spojitě prodloužení řádu 4 a 5 pro klasickou metodu RKF45. V tomto případě metoda řádu 5 také nesplňuje podmínku spojitosti (4.1.4) v uzlových bodech vzhledem k diskrétní metodě.

### 4.3 Aplikace odhadu lokální chyby

Je známo, že existují dva hlavní přístupy pro implementaci variabilního kroku diskrétní RK metody řádu  $p$ , s cílem na sestavení numerického řešení s globální chybou

$$\max_{1 \leq n \leq N} \|y(t_n) - y_n\|,$$

kteřá proporcionálně odpovídá uživatelskému parametru  $TOL$ , nazývaným *tolerance*. Koncept *proporcionální tolerance* byl přesně definován a prozkoumán Stetterem [33],[34].

První přístup spočívá v nalezení odhadu *lokální chyby na jednotku kroku*  $\tilde{\sigma}_{n+1}$

$$\sigma_{n+1} = \|z_{n+1}(t_{n+1} - y_{n+1})\|/h_{n+1}$$

v koncovém bodě  $t_{n+1}$  spolu s volbou kroku  $h_{n+1}$ , aby odhad lokální chyby  $\tilde{\epsilon}_{n+1}$  byl pod hranicí tolerance  $TOL$ .

Používáme dva hlavní způsoby na získání odhadu lokální chyby. Obecnější způsob spočívá v použití *Richardsonovy extrapolace* pro získání RK metody o jeden řád vyšší za účelem získání aproximace  $\tilde{y}_{n+1}$  v koncovém bodě  $t_{n+1}$ , která je použita pro stanovení

$$\tilde{\sigma}_{n+1} = \|\tilde{y}_{n+1} - y_{n+1}\|/h_{n+1}. \quad (4.3.1)$$

Tato technika je obvykle dost výpočetně náročná, potřebuje spočítat třikrát RK formuli pro každý integrační krok (v případě, že půlení kroku je použito pro extrapolaci). Nicméně v případě explicitních metod je možné se Richardsonově extrapolaci vyhnout, pokud máme vhodný pár metod, například *Rungeovy-Kuttovy-Fehlbergovy* metody (viz. Fehlberg [13]). Hlavní myšlenka použití vhodného páru metod spočívá v tom, že pomocí nějakých dalších stupňů je možné sestavit 2 různé aproximace  $y_{n+1}$  a  $\tilde{y}_{n+1}$  řádu  $p$  a  $p+1$  a použít je v (4.3.1).

Druhý přístup (vhodnější pro explicitní metody) spočívá v použití dvojice metod, kde se metoda nižšího řádu  $p-1$  použije pro získání aproximace  $\tilde{y}_{n+1}$ . Nejznámějším případem je Dormandova-Princova metoda [5]. Odhad

$$\tilde{\sigma}_{n+1} = \|\tilde{y}_{n+1} - y_{n+1}\|$$

lokální chyby  $\|z_{n+1}(t_{n+1}) - \tilde{y}_{n+1}\|$  pro metodu nižšího řádu je nahrazen odhadem  $\tilde{\sigma}_{n+1}$  lokální chyby na jednotku kroku a *lokální extrapolace* je použita pro zvolení kroku tak, aby  $\tilde{\sigma}_{n+1}$  byla pod hranicí tolerance  $TOL$ .

Krátké porovnání mezi těmito dvěma strategiemi ukáže, že druhý přístup je více efektivní, protože počítá další kroky pomocí přesnější metody. První strategie má ale obecně lepší proporcionální toleranci.

Možnost použít interpolant nám dává další alternativní strategie pro kontrolu délky kroku. Pro RK metody řádu  $p$  se spojitým prodloužením maximálního stejnoměrného řádu  $p$ , Enright [8],[9] navrhl monitorovat velikost *defektu*

$$\delta(t) = g(t, \eta(t)) - \eta'(t).$$

Pomocí Gröbnerovy-Alekseevovy nelineární variace konstant získáme:

$$z_{n+1}(t) - \eta(t) = \int_{t_n}^t K(t, x)[g(x, \eta(x)) - \eta'(x)]dx,$$

kde  $K(t, x)$  je vhodná variační matice závislá na interpolantu  $\eta(t)$  tak, že platí  $K(x, x) \equiv I$ , kde  $I$  je matice identity. Tedy pro dostatečně malé  $h_{n+1} \leq h_0$ , existuje konstanta  $K > 0$ , nezávislá na konkrétní RK metodě, tak že platí

$$\|z_{n+1}(t_{n+1}) - y_{n+1}\| \leq \max_{t_n \leq t \leq t_{n+1}} \|z_{n+1}(t) - \eta(t)\| \leq h_{n+1} K \max_{t_n \leq t \leq t_{n+1}} \|\delta(t)\|.$$

Tedy můžeme udělat závěr, že kontrolování defektu implikuje kontrolování lokální chyby na jednotku kroku  $\sigma_{n+1}$  a také dokonce stejnoměrné lokální chyby

$$\Sigma_{n+1} = \max_{t_n \leq t \leq t_{n+1}} \|z_{n+1}(t) - \eta(t)\|.$$

Víme, že není efektivní se pokusit počítat přesně velikost

$$\max_{t_n \leq t \leq t_{n+1}} \|\delta(t)\|. \quad (4.3.2)$$

Nicméně za vhodných podmínek pro interpolant  $\eta(t)$  existuje problém nezávislý na  $\theta^* \in [0, 1]$  takový, že  $\|\delta(t_n + \theta^* h_{n+1})\|$  je dobrá aproximace (4.3.2). Proto tedy volíme krok tak, aby  $\|\delta(t_n + \theta^* h_{n+1})\|$  bylo pod hranicí tolerance  $TOL$ .

Pro hlubší analýzu předchozích přístupů se odkazujeme na Enrighta [10].

Jako poslední strategii uvažujeme odhad lokální chyby od Zennara [35], [36]. Ten používá spojitého prodloužení  $\eta(t)$  maximálního stejnoměrného řádu  $p$ . Aproximaci  $\tilde{y}_{n+1}$  vyššího řádu  $p + 1$  můžeme získat pomocí aplikace kvadraturní formule na integrál

$$y_n + \int_{t_n}^{t_{n+1}} g(x, \eta(x)) dx.$$

Aproximace vyššího řádu má tedy formu:

$$\tilde{y}_{n+1} = y_n + h_{n+1} \sum_{i=1}^k \beta_i g(t_n + \pi_i h_{n+1}, \eta(t_n + \pi_i h_{n+1})), \quad (4.3.3)$$

pro vhodné abscisy  $\pi_1, \dots, \pi_k$ . Za účelem minimalizovat počet stupňů je vhodné použít superkonvergentní kvadraturní formuli, jako např. Gaussovy formule.

Zmíníme také, že Bellen a Zennaro [2] navrhli pro implicitní metody modifikaci kvadraturní formule (4.3.3) na implicitní formu:

$$\tilde{y}_{n+1} = y_n + h_{n+1} \sum_{i=1}^{k-1} \beta_i g(t_n + \pi_i h_{n+1}, \eta(t_n + \pi_i h_{n+1})) + h_{n+1} \beta_k g(t_{n+1}, \tilde{y}_{n+1}),$$

která je více vhodná pro získání aproximace s lepší stabilitou.

V každém případě, formuli (4.3.1) lze aplikovat a použít pro jakékoliv Rungeovy-Kuttovy-Fehlbergovy metody.

## 5 Rungeovy-Kuttovy metody pro zpožděné diferenciální rovnice

V této kapitole uvedeme standardní přístup, uvedený a analyzovaný v Kapitole 3, založený na aplikaci SRK metod z Kapitoly 4.

Připomeňme, že standardní přístup pro řešení ZDR s počáteční podmínkou

$$\begin{cases} y'(t) = f(t, y(t), y(t - \tau)), & t_0 \leq t \leq t_f, \\ y(t) = \phi(t), & t \leq t_0 \end{cases} \quad (5.0.1)$$

spočívá v řešení krok po kroku lokálních problémů:

$$\begin{cases} w'_{n+1}(t) = f(t, w_{n+1}(t), x(t - \tau)), & t_n \leq t \leq t_{n+1}, \\ w_{n+1}(t_n) = y_n, \end{cases}$$

kde

$$x(s) = \begin{cases} \phi(s) & \text{pro } s \leq t_0, \\ \eta(s) & \text{pro } t_0 \leq s \leq t_n, \\ w_{n+1}(s) & \text{pro } t_n \leq s \leq t_{n+1} \end{cases}$$

a  $\eta(s)$  je spojitá aproximace řešení, vypočtená metodou samotnou až do  $t_n$ . Spojitá RK metoda  $(A, b(\theta))$  s interpolantem první třídy určená pomocí (4.0.2), (4.1.1) nebo s interpolantem druhé třídy určená pomocí (4.0.2), (4.1.3) a (4.1.5), přejde pro řešení ZDR na (v  $Y$  notaci)

$$Y_{n+1}^i = y_n + h_{n+1} \sum_{j=1}^s a_{ij} f(t_{n+1}^j, Y_{n+1}^j, \eta(t_{n+1}^j - \tau)), \quad i = 1, \dots, s, \quad (5.0.2)$$

$$\eta(t_n + \theta h_{n+1}) = y_n + h_{n+1} \sum_{i=1}^s b_i(\theta) f(t_{n+1}^i, Y_{n+1}^i, \eta(t_{n+1}^i - \tau)), \quad 0 \leq \theta \leq 1, \quad (5.0.3)$$

kde  $t_{n+1}^i = t_n + c_i h_{n+1}$ ,  $c_i = \sum_{j=1}^s a_{ij}$ ,  $i = 1, \dots, s$ ,  $h_{n+1} = t_{n+1} - t_n$ . Tuto metodu budeme označovat jako *RK metoda pro ZDR* nebo ve zkratce jako *ZDR metoda*.

Stejně jako u RK metod pro ODR, pomocí

$$K_{n+1}^i = f(t_{n+1}^i, Y_{n+1}^i, \eta(t_{n+1}^i - \tau))$$

můžeme použít  $K$  notaci:

$$\eta(t_n + \theta h_{n+1}) = y_n + h_{n+1} \sum_{i=1}^s b_i(\theta) K_{n+1}^i, \quad 0 \leq \theta \leq 1, \quad (5.0.4)$$

$$K_{n+1}^i = f\left(t_{n+1}^i, y_n + h_{n+1} \sum_{j=1}^s a_{ij} K_{n+1}^j, \eta(t_{n+1}^i - \tau)\right), \quad i = 1, \dots, s. \quad (5.0.5)$$

Všimněme si, že pokud bychom použili RK metodu s abscisou  $c_i > 1$ , mohlo by to vést k situaci, kdy zpožděný argument  $t_{n+1}^i - \tau > t_{n+1}$  a nemohli bychom použít spojitě prodloužení. Proto budeme předpokládat, že RK metody pro ZDR (5.0.2), (5.0.3) jsou takové, že abscisy splňují  $c_i \in [0, 1]$ ,  $i = 1, \dots, s$ .

Následující věta vychází z obecné Věty 3.3.2.

**Věta 5.0.1** Uvažujme počáteční problém:

$$\begin{cases} y'(t) = f(t, y(t), y(t - \tau)), & t_0 \leq t \leq t_f, \\ y(t) = \phi(t), & t \leq t_0 \end{cases} \quad (5.0.6)$$

kde  $f(t, y, x)$  je funkce třídy  $C^p$  na intervalu  $[t_0, t_f] \times \mathbb{R}^d \times \mathbb{R}^d$ , počáteční funkce  $\phi(t)$  je třídy  $C^p$ . Dále předpokládejme, že síť  $\Delta = \{t_0, t_1, \dots, t_n, \dots, t_N = t_f\}$  obsahuje všechny body nespojitosti  $\xi_i = i\tau$ ,  $i = 1, \dots, p$ , ležící v  $[t_0, t_f]$ . Pokud má SRK metoda diskrétní řád  $p$  a stejnoměrný řád  $q$ , potom ZDR metoda (5.0.2),(5.0.3) má diskrétní globální řád  $p$  a stejnoměrný globální řád  $q' = \min\{p, q + 1\}$ , tedy

$$\max_{1 \leq n \leq N} \|y(t_n) - y_n\| = O(h^p)$$

a

$$\max_{t_0 \leq t \leq t_f} \|y(t) - \eta(t)\| = O(h^{q'}),$$

kde  $h = \max_{1 \leq n \leq N} h_n$ .

Poznamenejme tedy, že v případě použití metody s diskrétním řádem  $p$  a stejnoměrným řádem  $q < p - 1$ , ZDR metoda nemá zaručený diskrétní globální řád  $p$ .

Pokud máme metodu s konstantním krokem  $h = \tau/m$  pro nějaké přirozené číslo  $m \geq 1$ , potom zpožděný argument nabývá hodnot:

$$t_{n+1}^j - \tau = t_{n+1-m}^j = t_{n+1-m} + c_j h, \quad j = 1, \dots, s.$$

Pokud je SRK metoda přirozená, potom  $\eta(t_{n+1}^j) = Y_{n+1}^j$  a také  $\eta(t_{n+1}^j - \tau) = Y_{n+1-m}^j$  (viz. (4.1.8)). Tedy získáme diskrétní metodu

$$y_{n+1} = y_n + h \sum_{i=1}^s b_i f(t_{n+1}^i, Y_{n+1}^i, Y_{n+1-m}^i), \quad (5.0.7)$$

$$Y_{n+1}^i = y_n + h \sum_{j=1}^s a_{ij} f(t_{n+1}^j, Y_{n+1}^j, Y_{n+1-m}^j), \quad i = 1, \dots, s. \quad (5.0.8)$$

Tedy každá ZDR metoda (5.0.2), (5.0.3) založená na přirozené SRK metodě implementované s konstantním krokem  $h_{n+1} = h = \tau/m$ , kde  $m$  je přirozené číslo  $\geq 1$ , přejde na formuli (5.0.7),(5.0.8), která nepoužívá explicitně interpolaci. Takové metody budeme označovat jako *přirozené RK metody pro ZDR*.

Co se týče konvergence a řádu přesnosti (5.0.7), (5.0.8) můžeme použít obecné výsledky Věty 3.3.2. Na základě této věty, každá RK metoda pro ZDR má globální stejnoměrný řád  $q' = \min\{p, q + 1\}$ . Proto ZDR metoda zachovává diskrétní globální řád  $p$ , pokud má interpolant stejnoměrný řád  $q \geq p - 1$ . Obecně má ale interpolant přirozené SRK metody stejnoměrný řád  $q < p - 1$ . Přesto můžeme prokázat, že pro (5.0.7), (5.0.8) diskrétní globální řád  $p$  přirozené SRK metody je zachován bez ohledu na stejnoměrný řád  $q$  interpolantu.

**Věta 5.0.2** Pokud funkce  $f(t, y, x)$  je třídy  $C^p$  na intervalu  $[t_0, t_f] \times \mathbb{R}^d \times \mathbb{R}^d$ , potom přirozená RK metoda pro ZDR (5.0.7), (5.0.8) diskrétního řádu  $p$  má globální diskrétní řád  $p$ , tedy

$$\max_{1 \leq n \leq N} \|y(t_n) - y_n\| = O(h^p), \quad (5.0.9)$$

pro jakoukoliv počáteční funkci  $\phi(t)$  třídy  $C^p$  na intervalu  $[t_0 - \tau, t_0]$  a pro konstantní krok  $h = \tau/m$ , kde  $m$  je přirozené číslo  $\geq 1$ .

Důkaz lze nalézt v [3] (6.kapitola, strana 157).

Zejména podle Věty 5.0.2, superkovergentní přirozené SRK metody zachovávají superkonvergenci pro ZDR. Tedy dosažený řád v uzlových bodech je větší, než globální stejnoměrný řád. Je dobré poznamenat, že zachovávání superkonvergence je hlavně díky dvěma vlastnostem, které vyplývají z implementace přirozené SRK metody s konstantním krokem  $h = \tau/m$ , kde  $m$  je přirozené číslo  $\geq 1$

- (i) Zpožděný argument  $\alpha(t) = t - \tau$  zobrazí  $t$  z každého intervalu  $[t_n, t_{n+1}]$  do předchozího intervalu  $[t_{n-m}, t_{n+1-m}]$  a tedy v lokálním problému

$$\begin{cases} w'_{n+1}(t) = f(t, w_{n+1}(t), \eta(t - \tau)), & t_n \leq t \leq t_{n+1}, \\ w_{n+1}(t_n) = y_n \end{cases}$$

je funkce  $\eta(t - \tau)$  třídy  $C^\infty$ ;

- (ii) Pro každou abscissu  $c_i$  platí, že zpožděný argument  $\alpha(t) = t - \tau$  zobrazí každý bod  $t_{n+1}^i$  do bodu  $t_{n+1}^i - \tau = t_{n+1-m}^i$ , takže platí

$$\eta(t_{n+1}^i - \tau) = Y_{n+1-m}^i;$$

V případě libovolné volby uzlových bodů  $\{\xi_0 = t_0 < t_1 < \dots < t_m = \xi_1\}$  v prvním makro intervalu  $[\xi_0, \xi_1]$ , vlastnosti (i) a (ii) také platí v  $[t_0, t_f]$ , pokud uzlové body následujících makro intervalů  $[\xi_k, \xi_{k+1}]$ ,  $k = 1, 2, \dots$ , jsou definovány jako

$$t_{km+i} = t_i + k\tau, \quad i = 1, \dots, m, \quad k = 1, 2, \dots \quad (5.0.10)$$

Pokud síť  $\Delta$  splňuje podmínku (5.0.10), budeme ji označovat jako *omezenou síť*. Dále platí, že ZDR metoda založená na přirozené SRK metodě implementované na omezené síti  $\Delta$ , dostane tvar:

$$y_{n+1} = y_n + h_{n+1} \sum_{i=1}^s b_i f(t_{n+1}^i, Y_{n+1}^i, Y_{n+1-m}^i),$$

$$Y_{n+1}^i = y_n + h_{n+1} \sum_{j=1}^s a_{ij} f(t_{n+1}^j, Y_{n+1}^j, Y_{n+1-m}^j), \quad i = 1, \dots, s,$$

kde  $h_i = h_{km+i}$ ,  $i = 1, \dots, m$ ,  $k = 1, 2, \dots$  a tato metoda zachovává superkonvergenci v uzlových bodech.

## 6 Odhad lokální chyby a variabilní krok

Stejně jako u ODR metod, efektivní ZDR metoda by měla být implementována s možností variabilního kroku. Budeme uvažovat obecnou RK metodu pro ZDR v Y notaci (5.0.2), (5.0.3) nebo v K notaci (5.0.4),(5.0.5), pro které platí Věta 5.0.1.

Je známo, že strategie variabilního kroku má za cíl zajistit řešení s maximální globální chybou proporcionální k uživatelsky zadané toleranci (požadované přesnosti)  $TOL$  a zároveň minimalizovat výpočetní náročnost. V této kapitole budeme používat  $TOL$  ve smyslu *tolerance na jednotku kroku*.

V podkapitole 4.3 jsme uvedli nejvíce populární způsoby pro vytvoření odhadu lokální chyby pro ODR pomocí spojitých RK metod. Mimo případ měření velikosti defektu (viz. Enright a Hayashi [11]), mechanismus pro kontrolu délky kroku pro *hlavní metodu* řádu  $p$ , je nejčastěji založen na použití *pomocné metody* jiného řádu  $p'$ , obvykle  $p' = p + 1$  nebo  $p' = p - 1$ .

Rozšíření tohoto mechanismu pro ZDR není jednoduché a matematické zdůvodnění není očividné. Uvidíme, že správný mechanismus pro ZDR by měl kontrolovat zároveň diskrétní i stejnoměrnou lokální chybu. Tímto způsobem lze použít spojitou pomocnou metodu stejnoměrného řádu  $q' = q + 1$  nebo  $q' = q - 1$ , kde  $q$  je stejnoměrný řád hlavní metody.

Pro hlubší analýzu vztahu mezi lokální a globální chybou pro ZDR metody a požadavky na řád přesnosti interpolantu, odkazujeme čtenáře na Ooppelstrup [29], Filipi a Buchacker [14], Higham [20] a Guglielmi a Hairer [17].

### 6.1 Řád hlavní a pomocné metody

Předpokládejme, že hlavní spojitá RK metoda má diskrétní řád  $p$  a stejnoměrný řád  $q \geq p - 1$ , takže stejnoměrný globální řád je  $p$ .

Hlavní metoda řeší lokální problém pro ZDR

$$\begin{cases} w'_{n+1}(t) = f(t, w_{n+1}(t), x(t - \tau)) & t_n \leq t_{n+1}, \\ w_{n+1}(t) = y_n, \end{cases} \quad (6.1.1)$$

$$x(s) = \begin{cases} \phi(s) & \text{pro } s \leq t_0, \\ \eta(s) & \text{pro } t_0 \leq s \leq t_n, \\ w_{n+1}(s) & \text{pro } t_n \leq s \leq t_{n+1}, \end{cases}$$

kde pro  $s < t_n$  aproximace zpožděného řešení  $\eta(s)$  má stejnoměrný řád  $p$ , tedy

$$\max_{s \leq t_n} \|y(s) - \eta(s)\| = O(h^p), \quad (6.1.2)$$

kde  $h = \max_i h_i$ .

Nyní uvažujme pomocnou metodu řádu  $p'$ ,  $p' = p + 1$  nebo  $p' = p - 1$  pro řešení lokálního problému (6.1.1). Mechanismus pro kontrolu délky kroku vyžaduje, aby hlavní a pomocná metody dosahovaly svého řádu při řešení lokálního problému (6.1.1). Tedy chceme po aproximaci řešení hlavní metodou  $y_{n+1}$  a pomocnou metodou  $\tilde{y}_{n+1}$ , aby platilo

$$\|w_{n+1}(t_{n+1}) - y_{n+1}\| = O(h_{n+1}^{p+1}) \quad (6.1.3)$$

a

$$\|w_{n+1}(t_{n+1}) - \tilde{y}_{n+1}\| = O(h_{n+1}^{p'+1}), \quad (6.1.4)$$



pro fixní  $h_1, \dots, h_n$  a pro  $h_{n+1} \rightarrow 0$ . Pokud je funkce  $f$  z (6.1.1) třídy  $C^{p^*}$ , kde  $p^* = \max\{p, p'\}$ , tyto požadavky budou splněny. Obecně obraz zpožděného argumentu  $\alpha(t) = t - \tau$  současného intervalu  $[t_n, t_{n+1}]$  může obsahovat předchozí uzlové body, a proto interpolant  $\eta(t)$  je po částech polynomiální funkce, funkce  $\eta(t - \tau)$  může být pouze spojitá. Poznamenejme, že pokud je současný krok  $h_{n+1}$  dostatečně malý vzhledem k předchozím krokům, potom obraz  $\alpha([t_n, t_{n+1}])$  je celý obsažen uvnitř nějakého předchozího intervalu a tedy funkce  $\eta(t - \tau)$  je dostatečně hladká. Nicméně tento případ pravděpodobně nenastane, protože každý mechanismus pro kontrolu délky kroku poskytne sít splňující

$$\frac{\max_i h_i}{\min_i h_i} \leq K, \quad (6.1.5)$$

kde  $K$  je konstanta nezávislá na toleranci  $TOL$ . Proto musíme předpokládat, že lokální problém (6.1.1) má nehladkou pravou stranu. Nicméně podmínky (6.1.3) a (6.1.4) mohou být prokázány následujícím způsobem.

Uvažujme pomocný počáteční problém

$$\begin{cases} \tilde{w}'_{n+1}(t) = f(t, \tilde{w}_{n+1}(t), \tilde{x}(t - \tau)), & t_n \leq t \leq t_{n+1}, \\ \tilde{w}_{n+1}(t_n) = y_n, \end{cases} \quad (6.1.6)$$

$$x(s) = \begin{cases} \phi(s) & \text{pro } s \leq t_0, \\ \tilde{\eta}(s) & \text{pro } t_0 \leq s \leq t_n, \\ \tilde{w}_{n+1}(s) & \text{pro } t_n \leq s \leq t_{n+1}, \end{cases}$$

kde pro  $s < t_n$  nehladký interpolant  $\eta(s)$  je nahrazen funkcí  $\tilde{\eta}(s)$  třídy  $C^{p^*}$ . Dále předpokládejme, že funkce  $\tilde{\eta}(s)$  aproximuje  $\eta(s)$  do řádu  $p^*$ , tedy

$$\max_{s \leq t_n} \|\tilde{\eta}(s) - \eta(s)\| = O(h^{p^*}), \quad (6.1.7)$$

kde  $h = \max_i h_i$ . Všimněme si, že pomocí (6.1.5) přejde (6.1.7) na

$$\max_{s \leq t_n} \|\tilde{\eta}(s) - \eta(s)\| = O(h_{n+1}^{p^*}).$$

Řešení pomocného počátečního problému hlavní metodou označíme  $z_{n+1}$  a pomocnou metodou  $\tilde{z}_{n+1}$ . Protože je problém dostatečně hladký, dostáváme

$$\|\tilde{w}_{n+1}(t_{n+1}) - z_{n+1}\| = O(h_{n+1}^{p+1}) \quad (6.1.8)$$

a

$$\|\tilde{w}_{n+1}(t_{n+1}) - \tilde{z}_{n+1}\| = O(h_{n+1}^{p'+1}). \quad (6.1.9)$$

Dále pomocí (6.1.7) a Lipschitzovské spojitě závislosti na přesném a numerickém řešení, dostáváme

$$\|\tilde{w}_{n+1}(t_{n+1}) - w_{n+1}(t_{n+1})\| = O(h_{n+1}^{p^*+1}), \quad (6.1.10)$$

$$\|y_{n+1} - z_{n+1}\| = O(h_{n+1}^{p^*+1}) \quad (6.1.11)$$

a

$$\|\tilde{y}_{n+1} - \tilde{z}_{n+1}\| = O(h_{n+1}^{p^*+1}). \quad (6.1.12)$$

Celkově tedy máme

$$\begin{aligned} \|w_{n+1}(t_{n+1}) - y_{n+1}\| &\leq \|\tilde{w}_{n+1}(t_{n+1}) - w_{n+1}(t_{n+1})\| \\ &\quad + \|\tilde{w}_{n+1}(t_{n+1}) - z_{n+1}\| + \|z_{n+1} - y_{n+1}\| \end{aligned}$$

a s využitím (6.1.10), (6.1.8) a (6.1.11) dostaneme (6.1.3).

Analogicky

$$\begin{aligned} \|w_{n+1}(t_{n+1}) - \tilde{y}_{n+1}\| &\leq \|\tilde{w}_{n+1}(t_{n+1}) - w_{n+1}(t_{n+1})\| \\ &\quad + \|\tilde{w}_{n+1}(t_{n+1}) - \tilde{z}_{n+1}\| + \|\tilde{z}_{n+1} - \tilde{y}_{n+1}\| \end{aligned}$$

a užitím (6.1.10), (6.1.9) a (6.1.12) obdržíme (6.1.4).

Nastává tedy problém, jestli hladká funkce  $\eta(s)$  třídy  $C^{p^*}$ , použitá v pomocném problému (6.1.6), splňující podmínku (6.1.7), opravdu existuje.

### 6.1.1 Pomocná metoda řádu $p' = p - 1$

V tomto případě  $p^* = p$  a tedy pomocí (6.1.2), můžeme zvolit funkci  $\tilde{\eta}(s) = y(s)$ , která je hladká a splňuje (6.1.7). Pak je stejnoměrný řád  $q = p - 1$  interpolantu dostatečný.

### 6.1.2 Pomocná metoda řádu $p' = p + 1$

V tomto případě je  $p^* = p + 1$ , a tedy předpokládejme, že funkce  $f$  a  $\phi$  jsou třídy  $C^{p+1}$ . V tomto případě ale není lehké najít funkci  $\tilde{\eta}(s)$  splňující podmínku (6.1.7). Pomocí (6.1.2) můžeme vidět, že volba  $\tilde{\eta}(s) = y(s)$  nebude fungovat. Nicméně ukážeme, že v případě zvýšení stejnoměrného řádu hlavní metody na  $q = p$ , požadovaná funkce  $\tilde{\eta}(s)$  existuje.

Budeme pokračovat pomocí indukce skrz makro intervaly  $[\xi_{l-1}, \xi_l]$ ,  $l = 1, \dots, k$ , kde  $[\xi_{k-1}, \xi_k]$  obsahuje současný interval  $[t_n, t_{n+1}]$ .

Poznamenejme, že předpoklad (6.1.5) implikuje, že obraz zpožděného argumentu  $\alpha(t) = t - \tau$  každého intervalu  $[t_i, t_{i+1}]$  neobsahuje více než  $M$  předchozích intervalů, kde  $M$  je nezávislé na toleranci  $TOL$  a síti  $\Delta$ .

Proto pro každé  $l = k - 1, \dots, 1$ , obraz zpožděného argumentu současného intervalu  $[t_n, t_{n+1}]$ , který byl iterován  $(k - l)$ -krát, je obsažen mezi dvojicí uzlových bodů  $t_{n_l}$  a  $t_{n_l+m_l}$  spadajících do makro intervalu  $[\xi_{l-1}, \xi_l]$ , kde  $m_l \leq M^{k-l}$ . Nyní uvažujme řešení  $w_{n+1}(t)$  v prvním makro intervalu  $[\xi_0, \xi_1]$  lokálního pomocného problému

$$\begin{cases} \tilde{w}'_{n+1}(t) = f(t, \tilde{w}_{n+1}(t), \phi(t - \tau)), & t_{n_1} \leq t \leq t_{n_1+m_1}, \\ \tilde{w}_{n+1}(t_{n_1}) = y_{n_1}. \end{cases}$$

Toto je hladký problém a aproximace řešení  $\eta(t)$  má stejnoměrný řád  $q$ . Protože  $m_1 \leq M^{k-1}$ , dostáváme

$$\max_{t_{n_1} \leq t \leq t_{n_1+m_1}} \|\tilde{w}_{n+1}(t) - \eta(t)\| = O(h^{q+1})$$

jak pro  $q = p - 1$  nebo  $q = p$ .

Potom se posuneme do dalšího makro intervalu  $[\xi_1, \xi_2]$  a uvažujme lokální počáteční problém

$$\begin{cases} w'_{n_2+1}(t) = f(t, w_{n_2+1}(t), \eta(t - \tau)), & t_{n_2} \leq t \leq t_{n_2+m_2}, \\ w_{n_2+1}(t_{n_2}) = y_{n_2}, \end{cases}$$

jehož numerické řešení je také  $\eta(t)$ . Přítomnost  $\eta(t - \tau)$  ale činí problém nehladkým. Nahradíme ho tedy hladkým

$$\begin{cases} \tilde{w}'_{n_2+1}(t) = f(t, \tilde{w}_{n_2+1}(t), \tilde{w}_{n_1+1}(t - \tau)), & t_{n_2} \leq t \leq t_{n_2+m_2}, \\ \tilde{z}_{n_2+1}(t_{n_2}) = y_{n_2}, \end{cases} \quad (6.1.13)$$

kde  $\tilde{w}_{n_1+1}(t - \tau)$  je dobře definovaná funkce pro  $t \in [t_{n_2}, t_{n_2+m_2}]$ .

Protože  $m_2 \leq M^{k-2}$  můžeme ověřit:

$$\max_{t_{n_2} \leq t \leq t_{n_2+m_2}} \|\tilde{w}_{n_2+1}(t) - \eta(t)\| = O(h^{q+1}).$$

Toho můžeme dosáhnout pomocí numerického řešení (6.1.13) a trojúhelníkové nerovnosti za použití Lipschitzovy spojitě závislosti na přesném a numerickém řešení.

Nyní je jasné, že indukce platí. Požadovaná funkce  $\tilde{\eta}(t)$  třídy  $C^{p+1}$  splňující podmínku (6.1.7), je řešení  $\tilde{w}_{n_{k-1}+1}(t)$  hladkého počátečního problému

$$\begin{cases} \tilde{w}'_{n_{k-1}+1}(t) = f(t, \tilde{w}_{n_{k-1}+1}(t), \tilde{w}_{n_{k-2}+1}(t - \tau)), & t_{n_{k-1}} \leq t \leq t_{n_{k-1}+m_{k-1}}, \\ \tilde{w}_{n_{k-1}+1}(t_{n_{k-1}}) = y_{n_{k-1}}, \end{cases}$$

pokud  $q = p$ . Stejněměrný řád  $q$  je tedy nutno navýšit na  $q = p$ .

### 6.1.3 Spojitá pomocná metoda

Pro odhad stejnoměrné lokální chyby hlavní metody stejnoměrného řádu  $q$  můžeme použít spojitou pomocnou metodu stejnoměrného řádu  $q' \neq q$ ,  $q' = q - 1$  nebo  $q' = q + 1$ . Analýza je podobná, ale důkaz je jednodušší. Volba  $\tilde{\eta}(s) = y(s)$  vyhovuje oběma hodnotám  $q'$  a tedy stejnoměrný řád  $q$  hlavní metody je dostatečný.

## 6.2 Šíření chyby a kontrola délky kroku

V této podkapitole budeme zkoumat šíření lokální diskretní a lokální stejnoměrné chyby v intervalu  $[t_0, t_f]$  za účelem formulace maximální globální chyby. Analýzou zjistíme, které testy pro odhady lokální chyby jsou vhodné za účelem získání globální chyby proporcionální s tolerancí  $TOL$ . Nyní použijeme elementární verzi Gronwallova lemmatu (viz. [18]):

**Lemma 6.2.1** *Pokud  $\delta(t)$ ,  $\gamma(t)$  a  $\beta(t)$  jsou spojitě reálné funkce na intervalu  $[a, b]$ ,  $\gamma(t)$  je neklesající,  $\beta(t) \geq 0$  a platí*

$$\delta(t) \leq \gamma(t) + \int_a^t \beta(x)\gamma(x)dx, \quad a \leq t \leq b,$$

potom

$$\delta(t) \leq \gamma(t) \exp\left(\int_a^t \beta(x)dx\right).$$

Mějme síť  $\Delta = \{t_0, t_1, \dots, t_N = t_f\}$  a definujme maximální diskretní a stejnoměrnou globální chybu

$$e_n = \max_{i \leq n} \|y(t_i) - y_i\| \quad a \quad E_n = \max_{t \leq t_n} \|y(t) - \eta(t)\|.$$

Potom pro řešení  $w_{n+1}(t)$  lokálního problému (6.1.1) dostáváme

$$w_{n+1}(t) - y(t) = y_n - y(t_n) + \int_{t_n}^t (f(s, w_{n+1}(s), x(t - \tau)) - f(s, y(s), y(s - \tau))) ds,$$

a proto

$$\begin{aligned} \|w_{n+1}(t) - y(t)\| &\leq \|y(t_n) - y_n\| + \int_{t_n}^t L \|w_{n+1}(s) - y(s)\| ds \\ &\quad + \int_{t_n}^t M \|x(t - \tau) - y(t - \tau)\| ds, \end{aligned}$$

kde  $L$  a  $M$  jsou Lipschitzovské konstanty funkce  $f$ . Potom díky Lemmatu 6.2.1 získáváme

$$\|w_{n+1}(t) - y(t)\| \leq \left( e_n + M \int_{t_n}^t \|x(t - \tau) - y(t - \tau)\| ds \right) e^{L(t-t_n)},$$

kde

$$\|x(t - \tau) - y(t - \tau)\| \leq E_n.$$

Proto dostáváme

$$\max_{t_n \leq t \leq t_{n+1}} \|w_{n+1}(t) - y(t)\| \leq \frac{e^{Lh_{n+1}}}{1 - Mh_{n+1}e^{Lh}} e_n + \frac{Mh_{n+1}e^{Lh_{n+1}}}{1 - Mhe^{Lh}} E_n, \quad (6.2.1)$$

kde  $h = \max_i h_i$  je dostatečně malé.

Poznamenejme, že kdykoliv to bude možné či nutné, tak budeme nahrazovat současný krok  $h_{n+1}$  krokem  $h$  za účelem získání finální formule (6.2.5) pro maximální globální chybu.

Omezení pro diskrétní a stejnoměrnou globální chybu v  $[t_n, t_{n+1}]$  jsou

$$\|y(t_{n+1}) - y_{n+1}\| \leq h_{n+1}\sigma_{n+1} + \|w_{n+1}(t_{n+1}) - y(t_{n+1})\|$$

a

$$\max_{t_n \leq t \leq t_{n+1}} \|y(t) - \eta(t)\| \leq \Sigma_{n+1} + \max_{t_n \leq t \leq t_{n+1}} \|w_{n+1}(t) - y(t)\|,$$

kde  $\sigma_{n+1}$  je diskrétní lokální chyba na jednotku kroku

$$\sigma_{n+1} = \|w_{n+1}(t_{n+1}) - y_{n+1}\|/h_{n+1}$$

a  $\Sigma_{n+1}$  je stejnoměrná lokální chyba

$$\Sigma_{n+1} = \max_{t_n \leq t \leq t_{n+1}} \|w_{n+1}(t) - \eta(t)\|.$$

Díky nerovnosti (6.2.1) dostaneme

$$\|y(t_{n+1}) - y_{n+1}\| \leq h_{n+1}\sigma_{\Delta} + \frac{e^{Lh_{n+1}}}{1 - Mh_{n+1}e^{Lh}} e_n + \frac{Mh_{n+1}e^{Lh_{n+1}}}{1 - Mhe^{Lh}} E_n, \quad (6.2.2)$$

kde  $\sigma_{\Delta} = \max_i \sigma_i$  a tedy získáváme rekurzi pro maximální globální diskrétní chybu

$$e_{n+1} \leq h_{n+1}\sigma_{\Delta} + \frac{e^{Lh_{n+1}}}{1 - Mh_{n+1}e^{Lh}} e_n + \frac{Mh_{n+1}e^{Lh_{n+1}}}{1 - Mhe^{Lh}} E_n. \quad (6.2.3)$$

Analogicky, díky nerovnosti (6.2.1) získáme

$$\|y(t) - \eta(t)\| \leq \Sigma_\Delta + \frac{e^{Lh}}{1 - Mhe^{Lh}}e_n + \frac{Mhe^{Lh}}{1 - Mhe^{Lh}}E_n,$$

kde  $\Sigma_\Delta = \max_i \Sigma_i$ . Dále monotónnost  $e_n$  a  $E_n$  implikuje

$$E_n \leq E_{n+1} \leq \Sigma_\Delta + \frac{e^{Lh}}{1 - Mhe^{Lh}}e_n + \frac{Mhe^{Lh}}{1 - Mhe^{Lh}}E_n,$$

a tedy pro maximální stejnoměrnou globální chybu platí

$$E_n \leq \frac{1 - Mhe^{Lh}}{1 - 2Mhe^{Lh}}\Sigma_\Delta + \frac{e^{Lh}}{1 - 2Mhe^{Lh}}e_n. \quad (6.2.4)$$

Pomocí (6.2.3) a (6.2.4) po úpravě získáme

$$e_{n+1} \leq e^{Q_\Delta h_{n+1}}e_n + h_{n+1}(\sigma_\Delta + R_\Delta \Sigma_\Delta),$$

kde

$$Q_\Delta = L + \frac{3 - 4Mhe^{Lh}}{1 - 2Mhe^{Lh}}Me^{Lh} \quad \text{a} \quad R_\Delta = \frac{Me^{Lh}}{1 - 2Mhe^{Lh}}.$$

Všimněme si, že platí  $\lim_{h \rightarrow 0} Q_\Delta = L + 3M$  a  $\lim_{h \rightarrow 0} R_\Delta = M$ , takže pro všechny sítě  $\Delta$  s dostatečně malým krokem  $h \leq h_0$ , dostáváme

$$e_{n+1} \leq e^{Qh_{n+1}}e_n + h_{n+1}(\sigma_\Delta + R\Sigma_\Delta),$$

kde  $Q$  a  $R$  jsou nezávislé na  $\Delta$ . Proto můžeme udělat závěr, že existuje konstanta  $K > 0$  nezávislá na  $\Delta$ , pro kterou maximální diskrétní globální řád splňuje nerovnost

$$e_n \leq K(\sigma_\Delta + R\Sigma_\Delta) \quad (6.2.5)$$

pro všechna  $n = 1, \dots, N$ .

Všimněme si, že použitím (6.2.5) v (6.2.4) obdržíme podobné omezení pro maximální stejnoměrnou globální chybu.

Požadovanou proporcionalitu mezi globální chybou a danou tolerancí  $TOL$  získáme, pokud odhady lokální diskrétní chyby  $\tilde{\sigma}_{n+1}$  a lokální stejnoměrné chyby  $\tilde{\Sigma}_{n+1}$  splňují podmínku

$$\tilde{\sigma}_{n+1} \leq TOL \quad \text{a} \quad \tilde{\Sigma}_{n+1} \leq TOL \quad \text{pro všechna } n. \quad (6.2.6)$$

**Poznámka 6.2.2** *Konstanty  $K$  a  $R$  v (6.2.5) jsme získali za podmínky  $h \leq h_0$ , kde  $h_0 > 0$  dostatečně malé, aby platilo  $1 - 2Mh_0e^{Lh_0} > 0$ . Toto může být příliš přísné omezení pro velikost kroku, obzvlášť pro velké hodnoty  $L$  a  $M$ .*

Nyní zkoumejme rozdíl mezi hlavní RK metodou se stejnoměrným řádem  $q = p - 1$  a  $q = p$ , vzhledem k maximální globální chybě (6.2.5). Díky tomu budeme schopni formulovat podmínky, za kterých nebudeme muset v testu (6.2.6) zkoumat  $\tilde{\Sigma}_{n+1}$ . V části 6.1 jsme mohli pozorovat, že pro oba případy  $q = p - 1$  a  $q = p$ , diskrétní lokální chyba na jednotku kroku  $\sigma_{n+1}$  splňuje

$$\sigma_{n+1} = O(h_{n+1}^p). \quad (6.2.7)$$

Navíc podle úvahy v části 6.1 můžeme udělat závěr, že pouze v případě  $q = p$  platí

$$\sigma_{n+1} = \alpha_n h_{n+1}^p + O(h_{n+1}^{p+1}), \quad (6.2.8)$$

kde  $\alpha_n$  spojitě závisí na  $t_n, y_n$  a  $\eta(t)$  pro  $t \leq t_n$ , ale je nezávislá na  $h_{n+1}$ .

### 6.2.1 Stejnoměrný řád $q = p$

Pro stejnoměrnou lokální chybu  $\Sigma_{n+1}$  platí

$$\Sigma_{n+1} = \beta_n h_{n+1}^{p+1} + O(h_{n+1}^{p+2}), \quad (6.2.9)$$

kde  $\beta_n$  spojitě závisí na  $t_n, y_n$  a  $\eta(t)$  pro  $t \leq t_n$ , ale je nezávislá na  $h_{n+1}$ . Proto díky (6.2.8) dostáváme

$$\Sigma_{n+1} = \frac{\beta_n + O(h_{n+1})}{\alpha_n + O(h_{n+1})} h_{n+1} \sigma_{n+1}.$$

Dále necht existuje  $\kappa > 0$  taková, že platí

$$\frac{\beta_n}{\alpha_n} \leq \kappa,$$

pro dostatečně malé  $h = \max_i h_i$  (tj. pro  $y_n$  a  $\eta(t)$  dostatečně blízké s  $y(t_n)$  a  $y(t)$ ) a stejnoměrně vzhledem k  $t_n \in [t_0, t_f]$ . Lokální extrapolace nám pak dává

$$\Sigma_\Delta = hO(\sigma_\Delta). \quad (6.2.10)$$

Závěrem tedy můžeme říct, že pokud opakovaným zmenšením tolerance  $TOL$  se krok  $h$  bude čím dál tím víc zmenšovat a vliv stejnoměrné lokální chyby  $\Sigma_\Delta$  v (6.2.5) bude méně významný. Proto maximální globální chyba je určena hlavně diskrétní lokální chybou na jednotku kroku  $\sigma_\Delta$ . V takovéto situaci je rozumné použít podmínku

$$\tilde{\sigma}_{n+1} \leq TOL \quad \forall n \quad (6.2.11)$$

namísto (6.2.6). V případě, kdy stejnoměrná lokální chyba  $\Sigma_\Delta$  není zanedbatelná vzhledem k  $\sigma_\Delta$ , je vhodnější použít test (6.2.6).

### 6.2.2 Stejnoměrný lokální řád $q = p - 1$

Pro stejnoměrnou lokální chybu  $\Sigma_{n+1}$  platí:

$$\Sigma_{n+1} = \beta_n h_{n+1}^p + O(h_{n+1}^{p+1}),$$

kde  $\beta_n$  spojitě závisí na  $t_n, y_n$  a  $\eta(t)$  pro  $t \leq t_n$ , ale je nezávislá na  $h_{n+1}$ .

V tomto případě řád  $\Sigma_{n+1}$  se rovná řádu  $\sigma_{n+1}$ , a tedy místo (6.2.10) máme

$$\Sigma_\Delta = O(\sigma_\Delta). \quad (6.2.12)$$

Obecně ale neplatí (6.2.8), proto můžeme použít pouze (6.2.7). Tedy ověření (6.2.12) není tak zřejmé jako v předchozím případě  $q = p$ .

Vliv stejnoměrné lokální chyby  $\Sigma_\Delta$  a diskrétní lokální chyby na jednotku kroku  $\sigma_\Delta$  v (6.2.5) zůstává srovnatelný i v případě zmenšování tolerance  $TOL$ . Proto se doporučuje ověřit podmínky (6.2.6).

### 6.2.3 Shrnutí

Výše uvedená analýza je konzistentní s relevantními výsledky dostupnými v literatuře. Zejména citujeme Ooppelstruppa [29] a Highama [21]. Ooppelstrupp dokázal, že asymptotická expanze globální chyby alespoň do řádu  $p + 1$  existuje pouze, pokud je stejnoměrný

řád  $q = p$ , a tedy je možné sestavit správný odhad globální chyby pomocí Richardsonovi extrapolace. Higham dokázal, že v případě kontroly pouze diskrétní lokální chyby je nutné a dostatečné, aby stejnoměrný řád  $q$  interpolace byl rovný  $p$ , za účelem zajištění proporcionality mezi chybou a tolerancí.

Dále poznamenejme, že za účelem ilustrace vlivu interpolačního postupu na globální chybu ZDR řešiče, Arndt [1] vytvořil ukázkou řešiče se stejnoměrným řádem  $q = p - 1$ , kde velikost globální chyby je určena pouze stejnoměrnou lokální chybou, zatímco diskrétní lokální chyba je vždy rovna nule. I přesto, že jeho výběr rovnice je speciální, jeho příklad ukazuje, že v případě  $q = p - 1$ , může být velmi nespolehlivé kontrolovat pouze diskrétní lokální chybu.

### 6.3 Implementace

Nyní budeme ilustrovat výše uvedenou teorii vzorovou metodou a vhodným numerickým experimentem. Přesněji, budeme používat hlavní metodu stejnoměrného řádu  $q = p$  a pomocnou metodu diskrétního řádu  $p' = p + 1$ , získanou pomocí kvadraturní formule typu (4.3.3).

Při tomto nastavení aplikujeme postup popsáný v Podkapitole 6.2 pro explicitní RK metodu řádu  $p = 4$ , tzv. klasickou čtyř-stupňovou RK metodu

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

Uvažujme ZDR s konstantním zpožděním, řešíme tedy lokální počáteční problém v každém intervalu  $[t_n, t_{n+1}]$

$$\begin{cases} w'_{n+1}(t) = f(t, w_{n+1}(t), \eta(t - \tau)), & t_n \leq t \leq t_{n+1}, \\ w_{n+1}(t_n) = y_n. \end{cases} \quad (6.3.1)$$

Pomocí K notace, námi používaná RK metoda má předpis

$$\begin{aligned} K_{n+1}^1 &= f(t_n, y_n, \eta(t_n - \tau)), \\ K_{n+1}^2 &= f\left(t_n + \frac{1}{2}h_{n+1}, y_n + \frac{1}{2}h_{n+1}K_{n+1}^1, \eta\left(t_n + \frac{1}{2}h_{n+1} - \tau\right)\right), \\ K_{n+1}^3 &= f\left(t_n + \frac{1}{2}h_{n+1}, y_n + \frac{1}{2}h_{n+1}K_{n+1}^2, \eta\left(t_n + \frac{1}{2}h_{n+1} - \tau\right)\right), \\ K_{n+1}^4 &= f\left(t_n + h_{n+1}, y_n + h_{n+1}K_{n+1}^3, \eta\left(t_n + h_{n+1} - \tau\right)\right), \\ y_{n+1} &= y_n + h_{n+1} \left( \frac{1}{6}K_{n+1}^1 + \frac{1}{3}K_{n+1}^2 + \frac{1}{3}K_{n+1}^3 + \frac{1}{6}K_{n+1}^4 \right). \end{aligned} \quad (6.3.2)$$

Tato metoda má přirozené spojitě prodloužení řádu  $q = 3$  určené pomocí

$$\begin{aligned}
b_1(\theta) &= \left(\frac{2}{3}\theta^2 - \frac{3}{2}\theta + 1\right)\theta, \\
b_2(\theta) &= \left(-\frac{2}{3}\theta + 1\right)\theta^2, \\
b_3(\theta) &= \left(-\frac{2}{3} + 1\right)\theta^2, \\
b_4(\theta) &= \left(\frac{2}{3}\theta - \frac{1}{2}\right)\theta^2.
\end{aligned}$$

Nicméně pro naše účely potřebujeme interpolaci řádu  $q = p = 4$ . Pomocí postupu popsaném v [3] (5.kapitola, strana 135-140) získáme

$$\begin{aligned}
\eta_{(1)}(t_n + \theta h_{n+1}) &= y_n + h_{n+1} \left( \left( \frac{1}{6}d_2(\theta) + d_3(\theta) \right) K_{n+1}^1 + \frac{1}{3}d_2(\theta)K_{n+1}^2 \right. \\
&\quad \left. + \frac{1}{3}d_2(\theta)K_{n+1}^3 + \frac{1}{6}d_2(\theta)K_{n+1}^4 + d_4(\theta)K_{n+1}^5 + d_5(\theta)K_{n+1}^6 \right),
\end{aligned}$$

kde

$$\begin{aligned}
d_1(\theta) &= \frac{1}{2\theta_1-1}(\theta-1)^2(-3\theta^2+2(2\theta_1-1)\theta+2\theta_1-1), \\
d_2(\theta) &= \frac{1}{2\theta_1-1}\theta^2(3\theta^2-4(\theta_1+1)\theta+6\theta_1), \\
d_3(\theta) &= \frac{1}{2\theta_1(2\theta_1-1)}\theta(\theta-1)^2((1-3\theta_1)\theta+2\theta_1(2\theta_1-1)), \\
d_4(\theta) &= \frac{1}{2(\theta_1-1)(2\theta_1-1)}\theta^2(\theta-1)((2-3\theta_1)\theta+\theta_1(4\theta_1-3)), \\
d_5(\theta) &= \frac{1}{2\theta_1(2\theta_1-1)(\theta_1-1)}\theta^2(\theta-1)^2,
\end{aligned}$$

kde  $\theta_1$  je další abscisa,  $\theta_1 \in (0, 1)$ ,  $\theta_1 \neq \frac{1}{2}$ , a kde podpůrné stupně jsou

$$\begin{aligned}
K_{n+1}^5 &= K_{n+2}^1 = f(t_{n+1}, y_{n+1}, \eta(t_{n+1} - \tau)), \\
K_{n+1}^6 &= f(t_n + \theta_1 h_{n+1}, \eta_{(0)}(t_n + \theta_1 h_{n+1}, \eta(t_n + \theta_1 h_{n+1} - \tau))),
\end{aligned}$$

kde

$$\begin{aligned}
\eta_{(0)}(t_n + \theta h_{n+1}) &= y_n + h_{n+1} \left( \left( \frac{1}{6}\tilde{d}_2(\theta) + \tilde{d}_3(\theta)K_{n+1}^i + \frac{1}{3}\tilde{d}_2(\theta)K_{n+1}^2 \right) \right. \\
&\quad \left. + \frac{1}{3}\tilde{d}_2(\theta)K_{n+1}^3 + \frac{1}{6}\tilde{d}_2(\theta)K_{n+1}^4 + \tilde{d}_4(\theta)K_{n+1}^5 \right),
\end{aligned}$$



kde

$$\begin{aligned}\tilde{d}_1(\theta) &= (\theta - 1)^2(2\theta + 1), \\ \tilde{d}_2(\theta) &= \theta^2(3 - 2\theta), \\ \tilde{d}_3(\theta) &= \theta(\theta - 1)^2, \\ \tilde{d}_4(\theta) &= \theta^2(\theta - 1).\end{aligned}$$

Za účelem získání odhadu diskrétní lokální chyby potřebujeme pomocnou metodu vyššího řádu  $p' = p + 1 = 5$ , sestavíme tedy aproximaci  $\tilde{y}_{n+1}$  pomocí formule (4.3.3). Použijeme Lobattovy abscisy  $\pi_1 = \frac{5-\sqrt{5}}{10}$  a  $\pi_2 = \frac{5+\sqrt{5}}{10}$  a tedy další stupně jsou:

$$\begin{aligned}K_{n+1}^7 &= f(t_n + \pi_1 h_{n+1}, \eta_{(1)}(t_n + \pi_1 h_{n+1}), \eta(t_n + \pi_1 h_{n+1} - \tau)), \\ K_{n+1}^8 &= f(t_n + \pi_2 h_{n+1}, \eta_{(1)}(t_n + \pi_2 h_{n+1}), \eta(t_n + \pi_2 h_{n+1} - \tau)).\end{aligned}$$

Pak zavedeme

$$\tilde{y}_{n+1} = y_n + h_{n+1} \left( \frac{1}{12} K_{n+1}^1 + \frac{5}{12} K_{n+1}^7 + \frac{5}{12} K_{n+1}^8 + \frac{1}{12} K_{n+1}^5 \right). \quad (6.3.3)$$

Obvyklým způsobem (viz. (4.3.1)) odhadneme diskrétní lokální chybu na jednotku kroku

$$\tilde{\sigma}_{n+1} = \|\tilde{y}_{n+1} - y_{n+1}\|/h_{n+1}.$$

Aktuální krok je přijat pouze, pokud

$$\tilde{\sigma}_{n+1} \leq TOL, \quad (6.3.4)$$

kde  $TOL$  je uživatelsky zadaná tolerance.

Pokud je krok přijat, nový krok  $h_{\text{new}}$  je použit jako  $h_{n+2}$  v následujícím intervalu  $[t_{n+1}, t_{n+2}]$  a v opačném případě nahradí  $h_{n+1}$  v současném intervalu  $[t_n, t_{n+1}]$ . Nový krok spočteme pomocí formule

$$h_{\text{new}} = \max \left\{ \omega_{\min}, \min \left\{ \omega_{\max}, \rho \left( \frac{TOL}{\tilde{\sigma}_{n+1}} \right)^{\frac{1}{4}} \right\} \right\} \cdot h_{n+1}, \quad (6.3.5)$$

kde  $\omega_{\min} \in (0, 1)$ ,  $\omega_{\max} > 1$  a  $\rho \in (0, 1)$  jsou vhodné tzv. *bezpečnostní parametry*.

Uživatel může také zadat maximální počet zamítnutí na jeden krok a stanovit podmínku, jestli po zkrácení délky kroku nesmí dojít v následujícím kroku k zvětšení délky kroku.

Všimněme si, že pokud je splněna (6.3.4) a  $t_{n+1} + h_{\text{new}} > \xi_k$  pro bod nespojitosti  $\xi_k$  (který musí být obsažen v síti), zvolíme  $t_{n+1} = \xi_k$  a  $h_{n+2} = \xi_k - t_{n+1}$ .

Tuto metodu jsme implementovali v matlabu, lze ji nalézt v Příloze pod názvem “variabilni\_krok\_RK\_pro\_ZDR“, nedosahuje ale dobrých výsledků a potřebovala by určitou optimalizaci.

Protože stejnoměrný řád je  $q = p$ , kontrola pouze diskrétní lokální chyby je dostatečná. Nicméně uvedeme jako příklad možnost kontroly i stejnoměrné lokální chyby  $\Sigma_{n+1}$  za použití jiné pomocné metody stejnoměrného řádu  $q' = q - 1 = 3$ .

V tomto případě jde o vypočtení velikosti odhadu stejnoměrné lokální chyby:

$$\tilde{\Sigma}_{n+1} = \max_{0 \leq \theta \leq 1} \|\eta_{(0)}(t_n + \theta h_{n+1}) - \eta_{(1)}(t_n + \theta h_{n+1})\|. \quad (6.3.6)$$

V našem případě platí

$$\tilde{\Sigma}_{n+1} = \gamma_n h_{n+1}^4 + O(h_{n+1}^k),$$

kde  $\gamma_n$  spojitě závisí na  $t_n, y_n$  a  $\eta(t)$  pro  $t \leq t_n$ , ale nezávisí na  $h_{n+1}$ . Na druhou stranu, stejnoměrná lokální chyba  $\Sigma_{n+1}$  splňuje (6.2.9)

$$\Sigma_{n+1} = \beta_n h_{n+1}^5 + O(h_{n+1}^6),$$

kde  $\beta_n$  spojitě závisí na  $t_n, y_n$  a  $\eta(t)$  pro  $t \leq t_n$ , ale nezávisí na  $h_{n+1}$ . Proto dostáváme

$$\Sigma_{n+1} = \frac{\beta_n + O(h_{n+1})}{\gamma_n + O(h_{n+1})} h_{n+1} \tilde{\Sigma}_{n+1}.$$

Protože pro dostatečně malé  $h = \max_i h_i$  (tj. pro  $y_n$  a  $\eta(t)$  dostatečně blízko s  $y(t_n)$  a  $y(t)$ ) existuje  $\kappa > 0$  taková, že

$$\frac{\beta_n}{\gamma_n} \leq \kappa$$

stejněměrně vzhledem k  $t_n \in [t_0, t_f]$ , lokální extrapolace funguje a tedy můžeme udělat závěr, že platí

$$\Sigma_{\Delta} = O(TOL), \quad (6.3.7)$$

pokud platí

$$h_{n+1} \tilde{\Sigma}_{n+1} \leq TOL. \quad (6.3.8)$$

Poznamenejme, že díky (6.2.5) vztah (6.3.7) je nutný pro zachování proporcionality mezi chybou a tolerancí.

Test (6.3.8) uděláme pokaždé, když platí (6.3.4). Pokud (6.3.8) neplatí, upravíme další krok následující formulí

$$h_{\text{new2}} = \max \left\{ \omega_{\min}, \rho \left( \frac{TOL}{h_{n+1} \tilde{\Sigma}_{n+1}} \right)^{\frac{1}{5}} \right\} \cdot h_{n+1} \quad (6.3.9)$$

a opakujeme výpočet v aktuálním intervalu  $[t_n, t_{n+1}]$  s krokem  $h_{n+1} = h_{\text{new2}}$ .

I v případě, kdy je splněna podmínka (6.3.4), doporučujeme použít navíc ještě formuli (6.3.9), tedy dohromady s (6.3.5) máme

$$h_{n+2} = \max \left\{ \omega_{\min}, \min \left\{ \omega_{\max}, \rho \left( \frac{TOL}{\tilde{\sigma}_{n+1}} \right)^{\frac{1}{4}}, \rho \left( \frac{TOL}{h_{n+1} \tilde{\Sigma}_{n+1}} \right)^{\frac{1}{5}} \right\} \right\} \cdot h_{n+1}.$$

Pro vypočtení  $\tilde{\Sigma}_{n+1}$  z (6.3.6) stanovíme

$$\delta(\theta) = \eta_{(0)}(t_n + \theta h_{n+1}) - \eta_{(1)}(t_n + \theta h_{n+1})$$

a získáme

$$\begin{aligned} \delta(\theta) = h_{n+1} \frac{\theta^2(\theta-1)^2}{2(2\theta_1-1)} & \left( \frac{2\theta_1-1}{\theta_1} K_{n+1}^1 - (2K_{n+1}^2 + 2K_{n+1}^3 + K_{n+1}^4) \right. \\ & \left. + \frac{3\theta_1-2}{\theta_1-1} K_{n+1}^5 + \frac{1}{\theta_1(\theta_1-1)} K_{n+1}^6 \right), \end{aligned}$$

takže nakonec dostáváme

$$\tilde{\Sigma}_{n+1} = \left\| \delta\left(\frac{1}{2}\right) \right\|.$$

## 7 Analýza stability ZDR

### 7.1 Obecný nelineární systém ZDR s konstantním zpožděním

Uvažujme počáteční problém pro systém ZDR

$$\begin{cases} y'(t) = f(t, y(t), y(t - \tau)), & t \geq t_0, \\ y(t) = \phi(t), & t \leq t_0, \end{cases} \quad (7.1.1)$$

kde  $f \in C^0([t_0, t_f] \times \mathbb{C}^d \times \mathbb{C}^d)$ .

K tomu uvažujme jiný systém

$$\begin{cases} \tilde{y}'(t) = f(t, \tilde{y}(t), \tilde{y}(t - \tau)), & t \geq t_0, \\ \tilde{y}(t) = \tilde{\phi}(t), & t \leq t_0 \end{cases} \quad (7.1.2)$$

se stejnou pravou stranou  $f(t, y, x)$ , ale s jinou počáteční funkcí  $\tilde{\phi}(t)$ . Definujme

$$\begin{aligned} \delta y(t) &= \tilde{y}(t) - y(t), \\ \delta \phi(t) &= \tilde{\phi}(t) - \phi(t). \end{aligned}$$

**Definice 7.1.1** Řekneme, že systém ZDR (7.1.1) je kontraktivní, pokud pro každou jinou počáteční funkci  $\tilde{\phi}(t)$ , diference  $\delta y(t)$  řešení (7.1.1) a (7.1.2) splňuje

$$\|\delta y(t)\| \leq \max_{x \leq t_0} \|\delta \phi(x)\|, \quad t \geq t_0. \quad (7.1.3)$$

**Definice 7.1.2** Řekneme, že systém ZDR (7.1.1) nebo, ekvivalentně, jeho řešení  $y(t)$  je asymptoticky stabilní, pokud pro každou počáteční funkci  $\tilde{\phi}(t)$  diference  $\delta y(t)$  řešení (7.1.1) a (7.1.2) splňuje

$$\lim_{t \rightarrow +\infty} \delta y(t) = 0. \quad (7.1.4)$$

### 7.2 Lineární skalární testovací rovnice

Uvažujme lineární skalární ZDR s počáteční podmínkou

$$\begin{cases} y'(t) = \lambda y(t) + \mu y(t - \tau), & t \geq t_0, \\ y(t) = \phi(t), & t \leq t_0. \end{cases} \quad (7.2.1)$$

Následující věty vycházejí z obecnějších vět, dokázaných v [3] (9.kapitola, strana 250-256).

**Věta 7.2.1** Mějme vnitřní součin  $\langle \cdot, \cdot \rangle$  na  $\mathbb{C}^d$  a odpovídající normu  $\|\cdot\|$ , uvažujme systém ZDR (7.2.1). Pokud

$$\Re(\lambda) + |\mu| \leq 0, \quad (7.2.2)$$

potom je systém (7.2.1) kontraktivní (tj. platí (7.1.3)) pro všechny počáteční funkce  $\phi(t)$ .

**Věta 7.2.2** Mějme systém ZDR (7.2.1). Pokud platí

$$\Re(\lambda) + |\mu| < 0, \quad (7.2.3)$$

potom je systém nejen kontraktivní ale i asymptoticky stabilní (tj. platí (7.1.4)) pro všechny počáteční funkce  $\phi(t)$  a míra konvergence v (7.1.4) je alespoň exponenciální (tj. alespoň jako  $e^{-\alpha(t-t_0)}$  pro nějaké  $\alpha > 0$ ).

Je dobré poznamenat, že pro lineární ZDR s konstantním zpožděním, lze studovat stabilitu pomocí kořenů *charakteristické rovnice*

$$\zeta - \lambda - \mu e^{-\tau\zeta} = 0. \quad (7.2.4)$$

Je známo (viz. El'sgol'ts a Norkin [7]), že tato rovnice má nekonečně mnoho kořenů  $\zeta_i$ , každý s určitou násobností  $m_i$ . Leží v komplexní polorovině  $\Re(\lambda) < \alpha$  pro nějaké reálné číslo  $\alpha$  a jejich reálné části se kumulují v  $-\infty$ .

Řešení (7.2.1) má tvar v podobě nekonečné řady

$$y(t) = \sum_{i=1}^{\infty} \sum_{n_i=0}^{m_i-1} \alpha_{i,n_i} t^{n_i} e^{\zeta_i t}, \quad (7.2.5)$$

kde koeficienty  $\alpha_{i,n_i}$  jsou určeny počáteční funkcí  $\phi(t)$ . Z této reprezentace vyplývá, že nutná a postačující podmínka asymptotické stability (7.2.1) je, že všechny kořeny  $\zeta_i$  z (7.2.4) jsou takové, že  $\Re(\zeta_i) < 0$ .

Tato podmínka je splněna, pokud platí (7.2.3).

**Definice 7.2.3** *Oblast asymptotické stability  $S_\tau$  rovnice (7.2.1) je množina dvojic  $(\lambda, \mu)$  takových, že odpovídající řešení je asymptoticky stabilní pro každou počáteční funkci  $\phi(t)$ .*

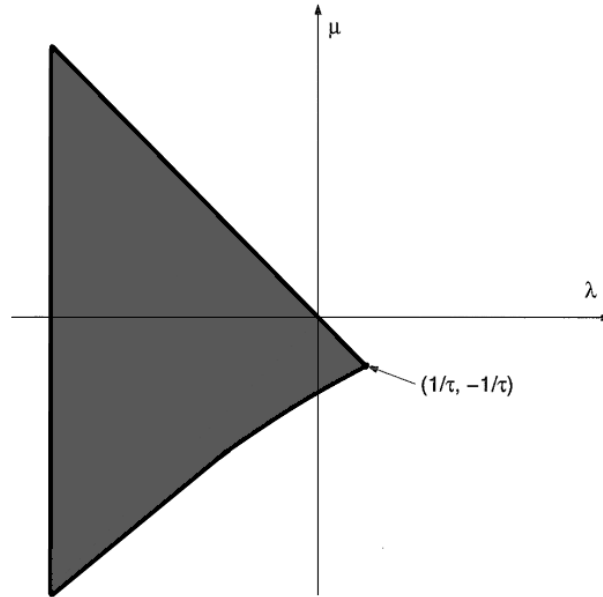
Efektivní technika pro nalezení oblasti asymptotické stability  $S_\tau$  je založena na metodě *D-dělení*, kterou představil Naimark [27] a která je často používaná v ruské literatuře (viz. El'sgol'ts a Norkin [7] a Kolmanovskii a Nosov [24]). Analýza oblasti asymptotické stability  $S_\tau$  pro reálné a komplexní  $\lambda, \mu$  byla značně prozkoumána Guglielmim [15], Guglielmim a Hairerem [16] a Masetem [25].

### Popis oblasti asymptotické stability pro reálné koeficienty

Uvažujme reálné koeficienty  $\lambda$  a  $\mu$ . Hlubší analýza charakteristické rovnice (7.2.4) ukáže, že oblast stability je větší než pouze  $\lambda + |\mu| < 0$ . Oblast asymptotické stability  $S_\tau$  je dána množinou dvojic  $(\lambda, \mu)$ , pro které platí (viz. Hayes [26])

$$\lambda < -\mu \quad \text{a} \quad \sqrt{\mu^2 - \lambda^2} < \frac{1}{\pi} \arccos(-\lambda/\mu).$$

Na obrázku 7.1 je vykreslena oblast asymptotické stability pro reálné  $\lambda, \mu$ .



Obrázek 7.1: Oblast asymptotické stability pro rovnici (7.2.1) s reálnými  $\lambda, \mu$

Počáteční problém pro ZDR

$$\begin{cases} y'(t) = \frac{1}{2}y(t) - y(t-1), & t \geq 0, \\ y(t) = \phi(t) = t + 1, & -1 \leq t \leq 0, \end{cases} \quad (7.2.6)$$

je příklad rovnice, kde řešení je asymptoticky stabilní, i když není splněna podmínka (7.2.3). Tedy podmínka (7.2.3) není nutnou podmínkou asymptotické stability. Nutná podmínka asymptotické stability tedy je

$$\lambda \leq \mu < -\lambda \quad (7.2.7)$$

Pro hlubší analýzu stability ZDR a numerických metod pro ZDR odkazujeme čtenáře na Jiaoxuna a Kuanga [23].

## 8 Příklady

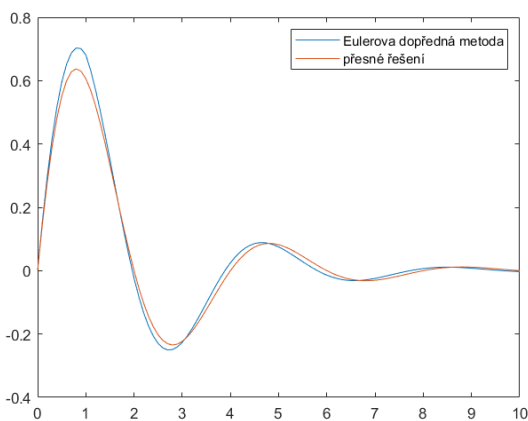
### Příklad 8.0.1 (Testovací rovnice)

Za účelem otestování možné ztráty přesnosti pro jednotlivé numerické metody, uvažujme testovací rovnici

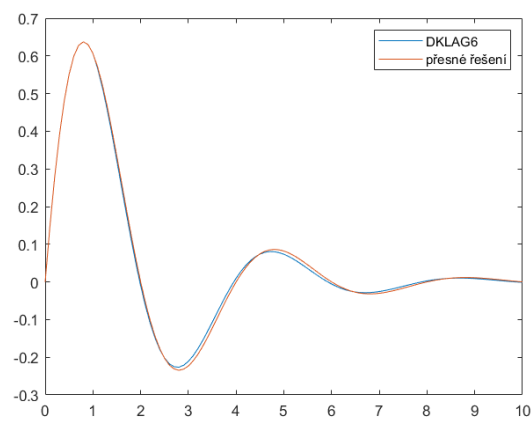
$$\begin{cases} y'(t) = ay(t) - \frac{\pi}{2}e^a y(t-1), & t \geq 0, \\ y(t) = \phi(t) = e^{at} \sin(\frac{\pi}{2}t), & t \leq 0, \end{cases} \quad (8.0.1)$$

kde  $a \in \mathbb{R}$  je parametr a přesné řešení  $y(t) = e^{at} \sin(\frac{\pi}{2}t)$  je funkce třídy  $C^\infty$  v  $[-1, +\infty)$ .

Nyní uvedeme numerické řešení (8.0.1), s parametrem  $a = -0.5$ , pomocí jednotlivých numerických metod s konstantním krokem (pro řešení na intervalu  $[0, 10]$  použijeme 100 kroků, resp. 10 kroků na jeden makro interval) a matlabovskou funkcí `dde23` (s nastavenou tolerancí  $TOL = 10^{-6}$ ).

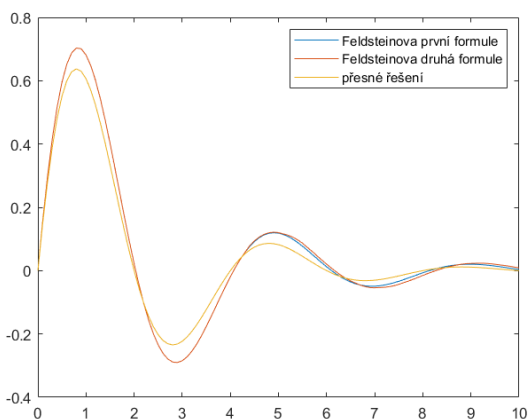


(a) pomocí Eulerovy dopředné metody

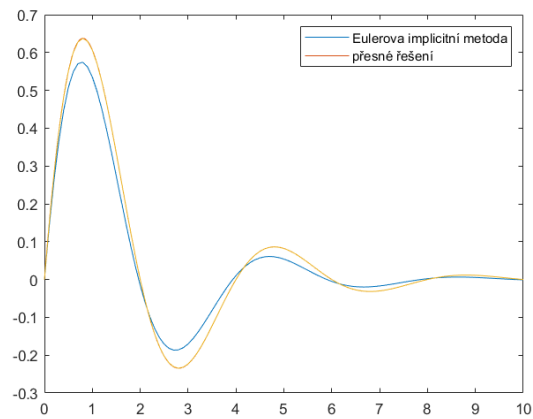


(b) pomocí metody DKL6

Obrázek 8.1: Numerické řešení (8.0.1) pro  $a = -0.5$  pomocí Bellmanovy metody kroků s Eulerovou dopřednou metodou a metody DKL6 ve srovnání s přesným řešením (8.0.1)

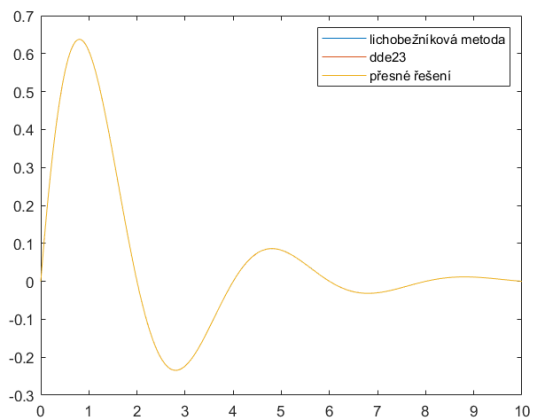


(a) pomocí Feldsteinovy metody

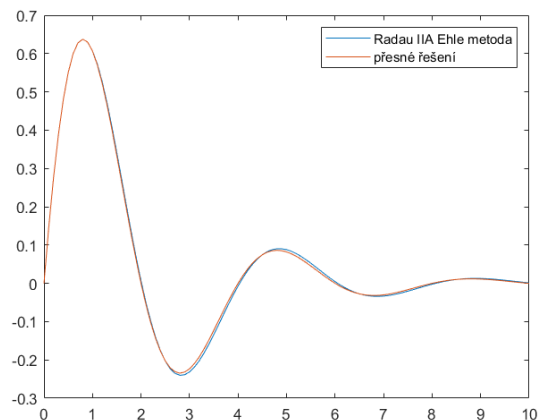


(b) pomocí Eulerovy implicitní metody

Obrázek 8.2: Numerické řešení (8.0.1) pro  $a = -0.5$  pomocí Feldsteinovy metody a metody kroků s Eulerovou implicitní metodou ve srovnání s přesným řešením (8.0.1)

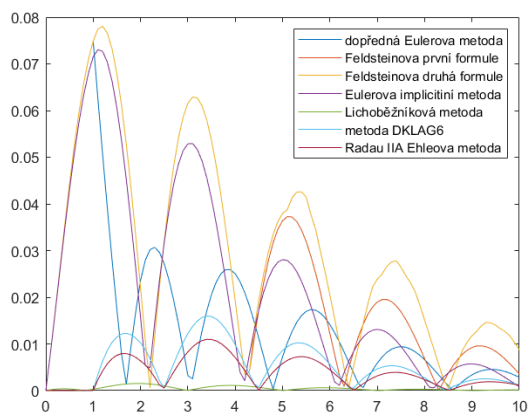


(a) pomocí lichoběžníkové metody a dde23

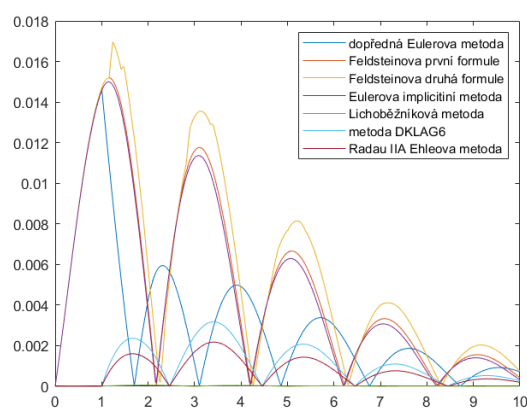


(b) pomocí Radau IIA Ehleovy metody

Obrázek 8.3: Numerické řešení (8.0.1) pro  $a = -0.5$  pomocí Radau IIA Ehleovy metody, metody kroků s lichoběžníkovou metodou a dde23 ve srovnání s přesným řešením (8.0.1)

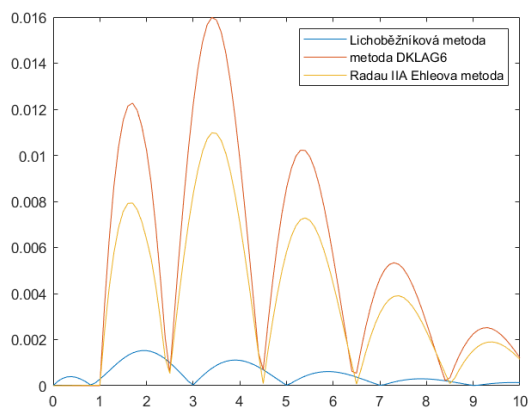


(a) Metody implementované se 100 kroky na intervalu  $[0, 10]$

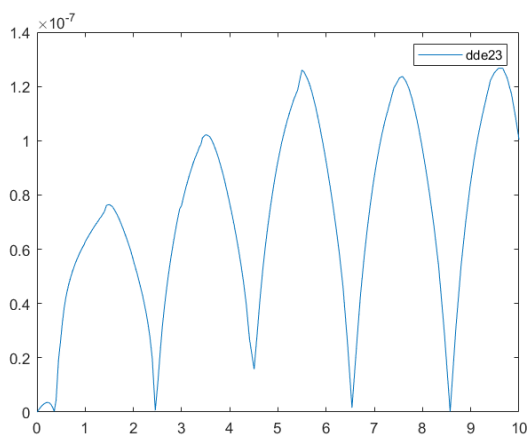


(b) Metody implementované s 500 kroky na intervalu  $[0, 10]$

Obrázek 8.4: Vizualizace absolutní chyby numerického řešení (8.0.1) pro uvedené numerické metody.



(a) Metoda DKLAGE, lichoběžníková metoda a Radau IIA Ehleova metoda se 100 kroky na intervalu  $[0, 10]$



(b) Metoda dde23

Obrázek 8.5: Vizualizace absolutní chyby numerického řešení (8.0.1) pro metodu DKLAGE, lichoběžníkovou metodu, Radau IIA Ehleovu metodu a matlabovskou funkcí dde23.

Můžeme pozorovat, že dobrých výsledků dosahovala DEKLAG6 metoda, Radau IIA Ehleova metoda (viz. Obrázky 8.1b, 8.3b a 8.4a) a jednoznačně nejlepších výsledků dosáhla lichoběžníková metoda (viz. Obrázky 8.3a a 8.5a), která je kolokační metoda s diskrétním a stejnoměrným řádem  $p = q = 2$ . Také si všimněme, že chyba metody dde23 je opravdu menší než zadaná tolerance  $TOL = 10^{-6}$  (viz. 8.5b). Pozorujme, že když jsme zvětšili počet kroků  $5\times$ , absolutní chyba se přibližně  $5\times$  zmenšila (viz. Obrázky 8.4a a 8.4b).

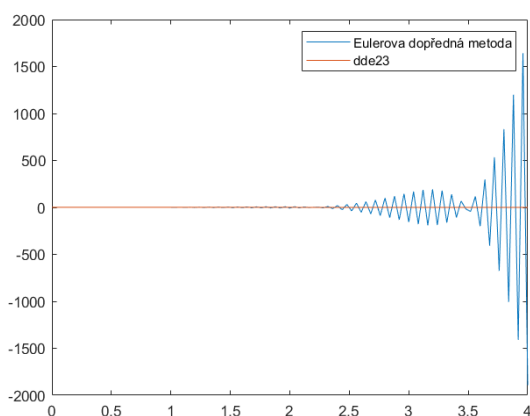


## Příklad 8.0.2 (Test stability)

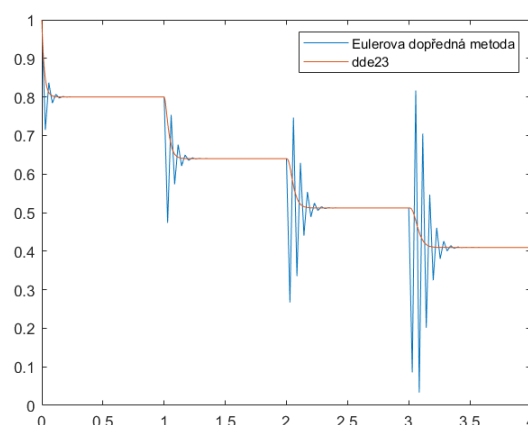
Uvažujme lineární rovnici

$$\begin{cases} y'(t) = \lambda y(t) - \frac{4}{5}\lambda y(t-1), & t \geq 0, \\ y(t) = 1, & t \leq 0, \end{cases} \quad (8.0.2)$$

kde  $\lambda \in \mathbb{R}$  je parametr. Dle (7.2.3) je tato rovnice pro  $\lambda < 0$  asymptoticky stabilní. Zvolme  $\lambda = -50$  a otestujme naše vytvořené numerické metody na (8.0.2) a srovnajme s dde23.

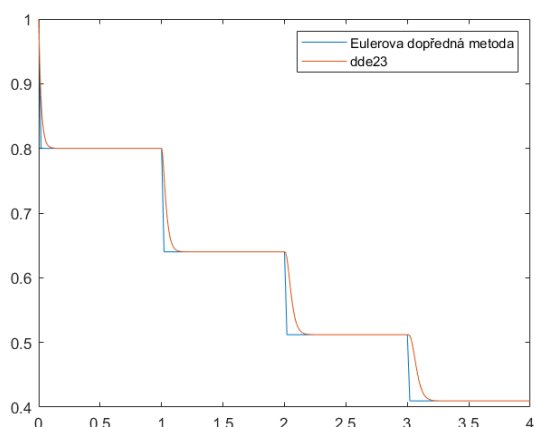


(a) Dopředná Eulerova metoda s 25 kroky na jeden makro interval

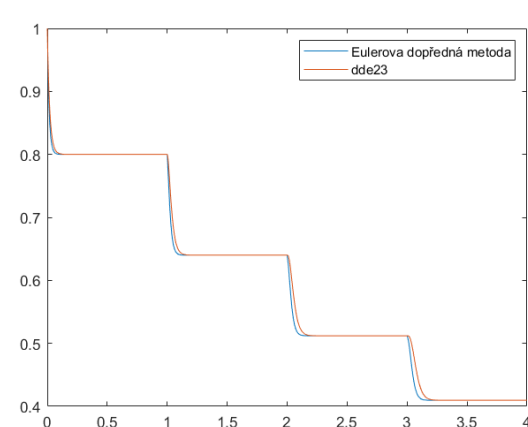


(b) Dopředná Eulerova metoda s 35 kroky na jeden makro interval

Obrázek 8.6: Numerické řešení (8.0.2) s  $\lambda = -50$  pomocí Bellmanovy metody kroků s dopřednou Eulerovou metodou v porovnání s řešením (8.0.2) pomocí dde23

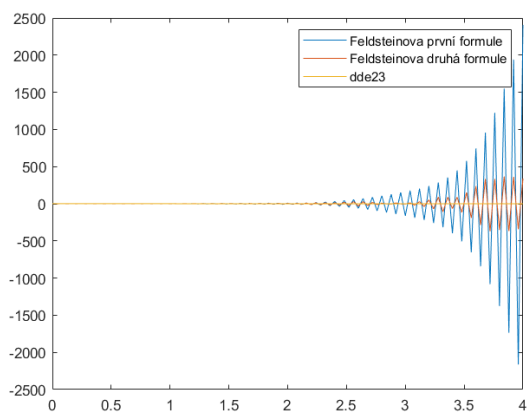


(a) Dopředná Eulerova metoda s 50 kroky na jeden makro interval

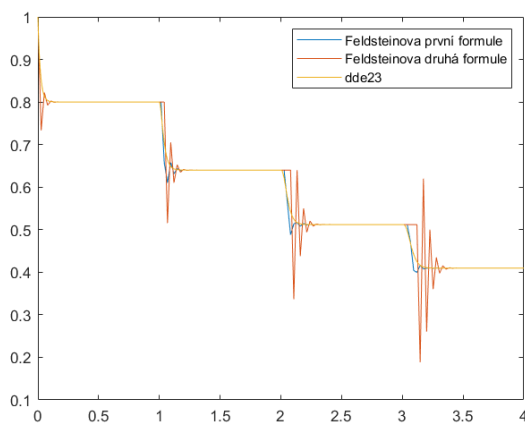


(b) Dopředná Eulerova metoda s 100 kroky na jeden makro interval

Obrázek 8.7: Numerické řešení (8.0.2) s  $\lambda = -50$  pomocí Bellmanovy metody kroků s dopřednou Eulerovou metodou v porovnání s řešením (8.0.2) pomocí dde23

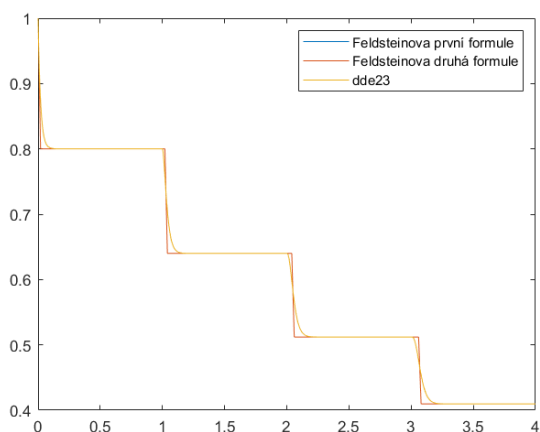


(a) Feldsteinova metoda s 100 kroky na intervalu  $[0, 4]$

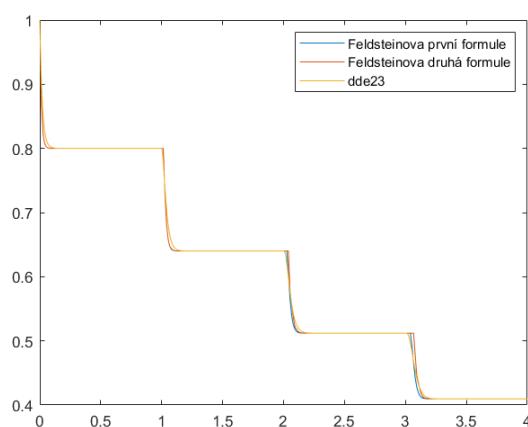


(b) Feldsteinova metoda s 150 kroky na intervalu  $[0, 4]$

Obrázek 8.8: Numerické řešení (8.0.2) s  $\lambda = -50$  pomocí Feldsteinovy metody v porovnání s řešením (8.0.2) pomocí dde23

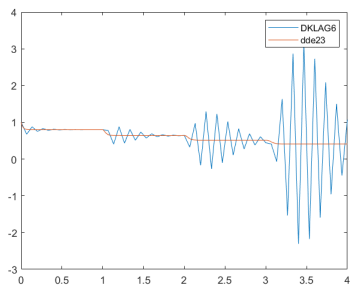


(a) Feldsteinova metoda s 200 kroky na intervalu  $[0, 4]$

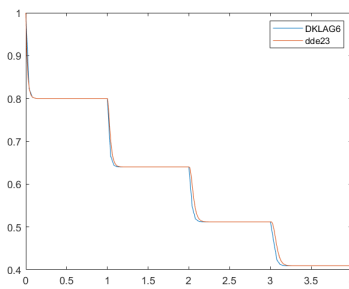


(b) Feldsteinova metoda s 300 kroky na intervalu  $[0, 4]$

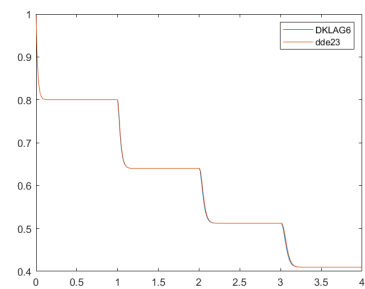
Obrázek 8.9: Numerické řešení (8.0.2) s  $\lambda = -50$  pomocí Feldsteinovy metody v porovnání s řešením (8.0.2) pomocí dde23



(a) Metoda DKLAG6 s 15 kroky na jeden makro interval

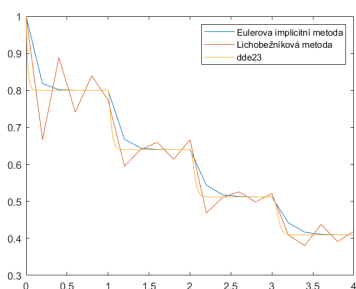


(b) Metoda DKLAG6 s 25 kroky na jeden makro interval

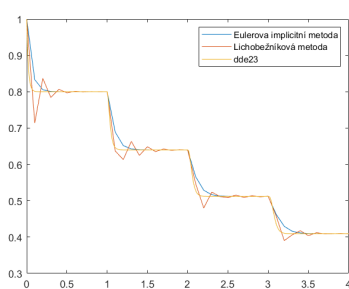


(c) Metoda DKLAG6 se 100 kroky na jeden makro interval

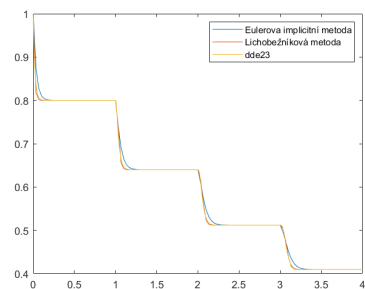
Obrázek 8.10: Numerické řešení (8.0.2) s  $\lambda = -50$  pomocí DKLAG6 metody v porovnání s řešením (8.0.2) pomocí dde23



(a) Eulerova implicitní metoda a lichoběžníková metoda s 5 kroky na jeden makro interval

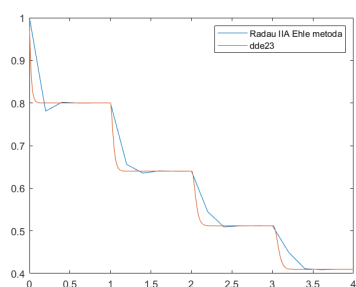


(b) Eulerova implicitní metoda a lichoběžníková metoda s 10 kroky na jeden makro interval

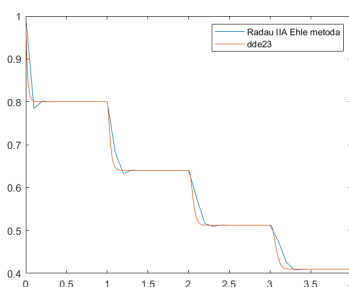


(c) Eulerova implicitní metoda a lichoběžníková metoda s 30 kroky na jeden makro interval

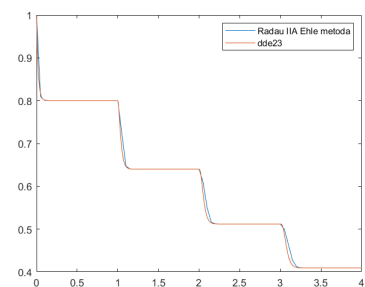
Obrázek 8.11: Numerické řešení (8.0.2) s  $\lambda = -50$  pomocí metody kroků s Eulerovou implicitní metodou a lichoběžníkovou metodou v porovnání s řešením (8.0.2) pomocí dde23



(a) Radau IIA Ehleova metoda s 5 kroky na jeden makro interval



(b) Radau IIA Ehleova metoda s 10 kroky na jeden makro interval



(c) Radau IIA Ehleova metoda s 20 kroky na jeden makro interval

Obrázek 8.12: Numerické řešení (8.0.2) s  $\lambda = -50$  pomocí Radau IIA Ehleovy metody v porovnání s řešením (8.0.2) pomocí dde23

Ilustrovali jsme tedy, že i když máme asymptoticky stabilní přesné řešení, některé numerické metody mohou být nestabilní (viz. Obrázky 8.6a, 8.6b, 8.8a, 8.8b a 8.1b).

Feldsteinova metoda není v tomto směru vhodná (viz. Obrázky 8.8a, 8.8b), protože její algoritmus nevyžaduje přítomnost všech bodů nespojitosti (1.1.2) v síti. Mohli jsme pozorovat, že Radau IIA Ehleova metoda, lichoběžníková metoda a implicitní Eulerova metoda byly stabilní i pro malé počty kroků na jeden makro interval (viz, Obrázky 8.11a a 8.12a), zatímco dopředná Eulerova metoda, Feldsteinova metoda a metoda DKL6 pro určitý počet kroků na jeden makro interval byly nestabilní (viz. Obrázky 8.6a, 8.6b, 8.8a, 8.8b a 8.1b). Pro počáteční problémy, kde by mohl nastat problém se stabilitou, doporučujeme volbu Radau IIA Ehleovy metody (viz. Obrázky 8.12a, 8.12b a 8.12c) nebo metody s variabilním krokem.

### Příklad 8.0.3 (Modifikovaný Verhulstův model)

Jedná se o jednoduchý biologický model růstu populace

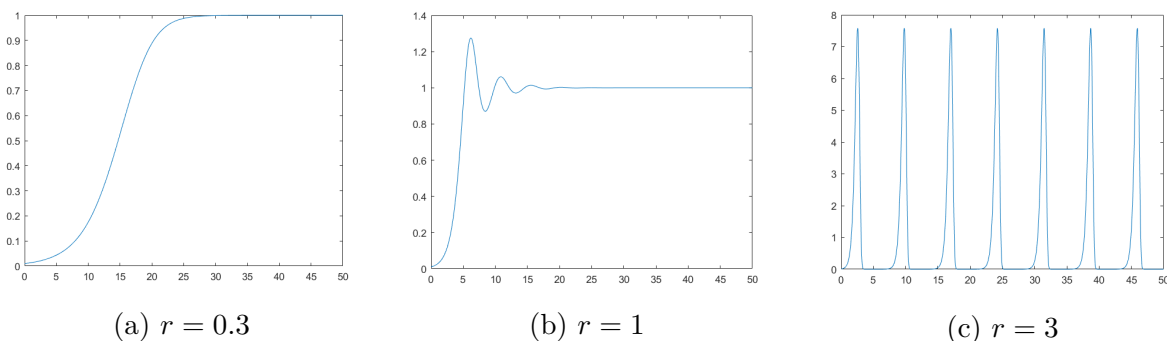
$$\begin{cases} y'(t) = ry(t) \left(1 - \frac{y(t-\tau)}{K}\right), & t \geq 0, \\ y(t) = \phi(t) & t \leq 0, \end{cases} \quad (8.0.3)$$

kde  $y(t)$  je stav populace,  $\tau$  představuje dobu dospívání,  $r > 0$  je koeficient růstu populace,  $K > 0$  je nosnost prostředí a  $\phi(t)$  je počáteční funkce (např. konstantní funkce). Pokud by  $\tau = 0$ , dostáváme klasický Verhulstův model, tedy logistickou rovnici.

Pro různá  $r$  mohou nastat celkově 3 typy řešení:

- pro  $r \in (0, \frac{1}{e})$   $y$  se monotónně přibližuje k  $y = K$  a má limitu  $\lim_{t \rightarrow +\infty} y(t) = K$  (viz. Obrázek 8.13a);
- pro  $r \in [\frac{1}{e}, \frac{\pi}{2})$   $y$  osciluje kolem hodnoty  $y = K$  a má limitu  $\lim_{t \rightarrow +\infty} y(t) = K$  (viz. Obrázek 8.13b);
- pro  $r \geq \frac{\pi}{2}$   $y$  osciluje a nemá limitu (viz. Obrázek 8.13c);

Uvedeme nyní numerické řešení pomocí matlabovského programu dde23 (s nastavenou tolerancí  $TOL = 10^{-6}$ ) pro  $K = 1$ ,  $\tau = 1$ ,  $\phi(t) = 0.01$  a pro jednotlivá  $r = 0.3$ ,  $r = 1$  a  $r = 3$ .

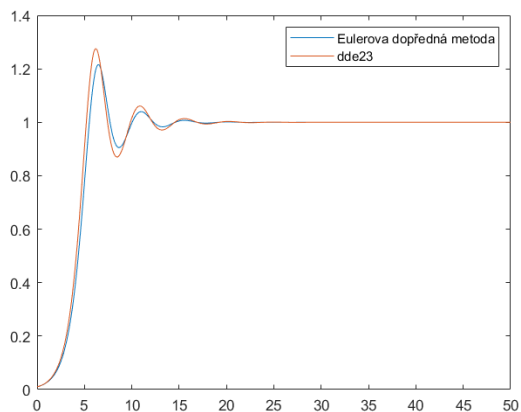


Obrázek 8.13: Numerické řešení (8.0.3) pomocí dde23 pro jednotlivá  $r$

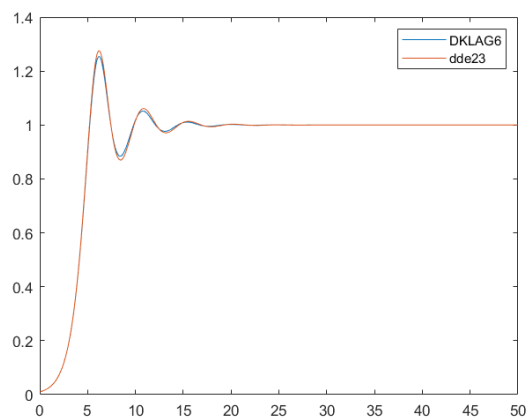
Poznamenejme, že pro případ  $r = 3$  (viz. Obrázek 8.13c), řešení po dosažení prvního vrcholu klesne až na hodnotu přibližně  $y = 10^{(-6)}$ , to by reálném případě znamenalo vymření dané populace.

Nyní použijeme vybrané metody (Feldsteinova metoda, Bellmanova metoda kroků, metoda kroků s Eulerovou implicitní metodou a lichoběžníkovou metodou, Radau IIA

Ehleova metoda a metoda DKLAG6) a pro  $r = 1$  je porovnáme s řešením pomocí dde23 (na interval  $[0, 50]$  použijeme 500 kroků, neboli 10 kroků na jeden makro interval).

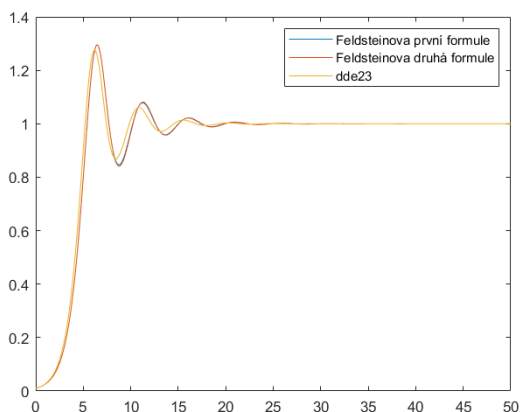


(a) pomocí Bellmanovy metody kroků

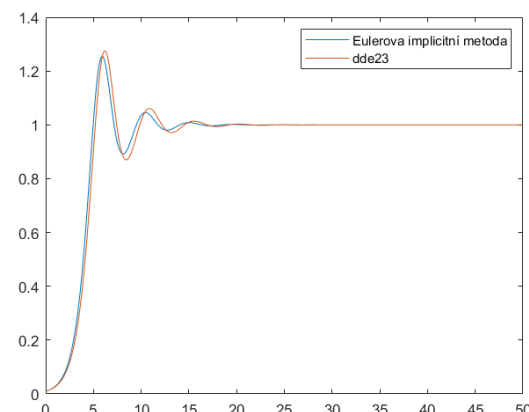


(b) pomocí metody DKLAG6

Obrázek 8.14: Numerické řešení (8.0.3) pro  $r = 1$  pomocí Bellmanovy metody kroků a metody DKLAG6 ve srovnání s metodou dde23

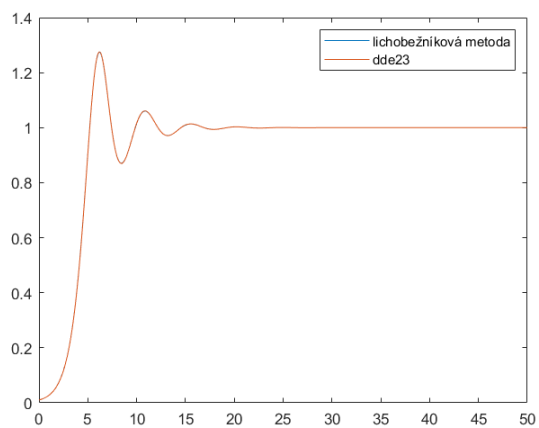


(a) pomocí Feldsteinovy metody

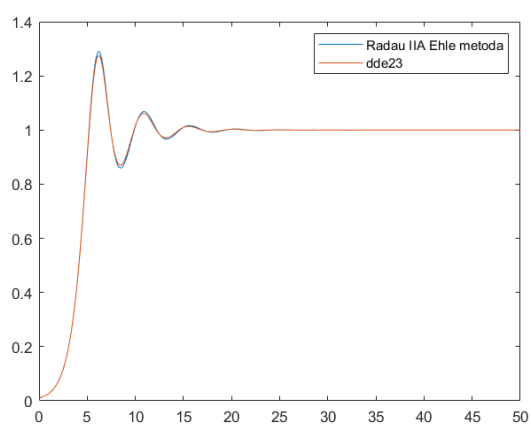


(b) pomocí implicitní eulerovy metody

Obrázek 8.15: Numerické řešení (8.0.3) pro  $r = 1$  pomocí Feldsteinovy metody a implicitní eulerovy metody ve srovnání s metodou dde23



(a) pomocí lichoběžníkové metody



(b) pomocí Radau IIA Ehle metody

Obrázek 8.16: Numerické řešení (8.0.3) pro  $r = 1$  pomocí lichoběžníkové metody a Radau IIA Ehle metody ve srovnání s metodou dde23

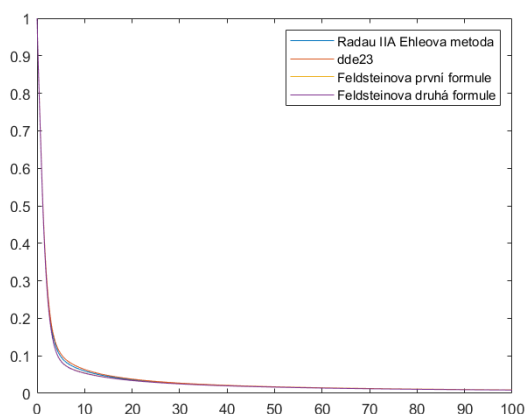
### Příklad 8.0.4 (Vývoj směnného kurzu)

Model představující vývoj směnného kurzu mezi 2 měnami, uvedený v [12]

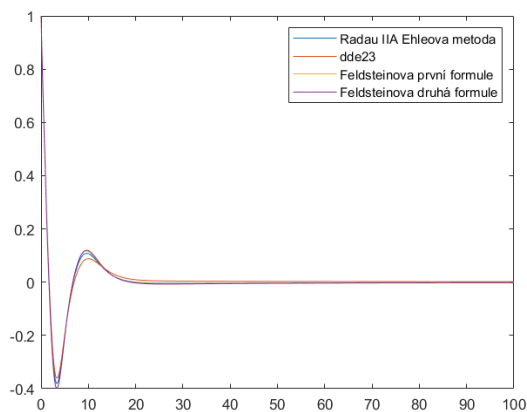
$$\begin{cases} y'(t) = a(y(t) - y(t-1) - |y(t)|y(t)), & t \geq 0, \\ y(t) = \phi(t), & t \leq t_0, \end{cases} \quad (8.0.4)$$

kde  $a > 0$  je parametr. Pokud se směnný kurz zvýší a platí  $y(t) > y(t-1)$ , spekulanti mohou předpokládat další růst a nakoupí cizí měnu. Poptávka po cizí měně stoupne a směnný kurz bude dále růst. Na druhou stranu v případě snížení směnného kurzu tak, že platí  $y(t) < y(t-1)$ , bude tendence cizí měnu prodat (než klesne na hodnotě ještě více), poptávka klesne a směnný kurz bude dále klesat. Po nějaké době se absolutní výchylka  $|y(t)|$  zvětší natolik, že spekulanti začnou předpokládat možnou změnu trendu a začnou prodávat/kupovat (tento jev ilustruje prvek  $-|y(t)|y(t)$ ). Znehodnocení (nebo zhodnocování) domácího kurzu vede k růstu absolutní výchylky  $|y(t)|$  a zároveň nelineární zpětná vazba  $-|y(t)|y(t)$  má vliv na snížení absolutní výchylky, proto očekáváme oscilační efekt.

Nyní uvedeme numerické řešení (8.0.4) s volbou  $\phi(t) = 1$  pomocí Feldsteinovy metody, Radau IIA Ehleovy metody (obě metody implementujeme s 10 kroky na jeden makro interval) a dde23 pro různá  $a$ .

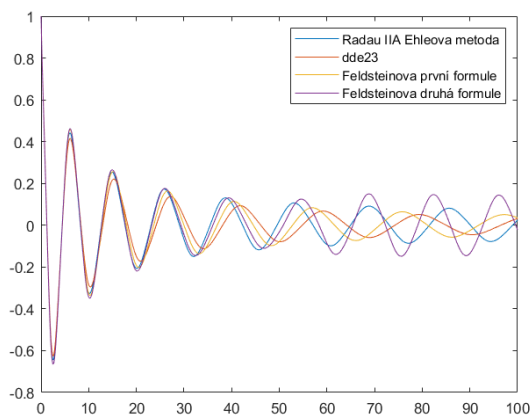


(a) numerické řešení (8.0.4) pro  $a = 0.5$

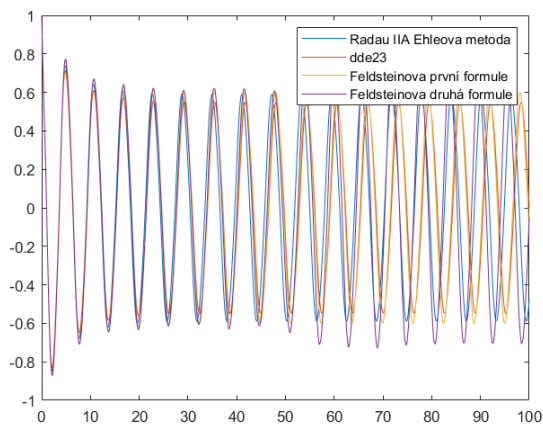


(b) numerické řešení (8.0.4) pro  $a = 0.8$

Obrázek 8.17: numerické řešení (8.0.4) pro konkrétní volby  $a$  pomocí Feldsteinovy metody, Radau IIA Ehleovy metody a dde23



(a) numerické řešení (8.0.4) pro  $a = 1$



(b) numerické řešení (8.0.4) pro  $a = 1.2$

Obrázek 8.18: numerické řešení (8.0.4) pro konkrétní volby  $a$  pomocí Feldsteinovy metody, Radau IIA Ehleovy metody a dde23

Můžeme tedy pozorovat, že pro  $a \leq 1$  řešení má tendenci se stabilizovat (viz. Obrázky 8.17a, 8.17b a 8.18a) a pro menší hodnoty  $a$  (např.  $a = 0.5$ ) nedochází k oscilaci (viz. Obrázek 8.17a, jedná se tedy o model systému, kde vliv spekulací není tak velký. Pro vyšší hodnoty  $a$  (např.  $a = 1.2$ ) dochází k oscilaci a kurz nebude stabilní díky silnému vlivu spekulací (viz. Obrázek 8.18b).



## Závěr

Zpožděný člen v ZDR s sebou přináší určité komplikace. Ilustrovali jsme, že aplikace jednokrokových numerických metod pro řešení ZDR s počáteční podmínkou se neobejde bez hlubší znalosti tzv. spojitých numerických metod. V této práci jsme tedy uvedli potřebné informace a ilustrovali aplikaci spojitých metod pro ODR na řešení ZDR s počáteční podmínkou. Oproti numerickým metodám pro ODR spočívá hlavní rozdíl v tom, že musíme implementovat numerickou metodu pro ZDR tak, že používá pouze uzlové body, kde zpoždění či nespojitost pravé strany určuje určité požadavky na tvar sítě, nebo musíme průběžně počítat a ukládat tzv. spojitě prodloužení numerické metody.

Na základě uvedené teorie jsme sestavili několik vybraných numerických metod pro řešení ZDR. Jejich implementaci v Matlabu lze nalézt v Příloze. Jejich aplikace a studium vlastností numerického řešení jsme uvedli v poslední kapitole na vybraných příkladech.

V této práci jsme uvedli základy jednokrokových numerických metod pro zpožděné rovnice s konstantním zpožděním. Hlavním zdrojem byla monografie [3], která má daleko širší záběr a mnohé výsledky jsou formulovány v obecném duchu. Restrikce úvah na jednokrokové metody a konstantní zpoždění vedla ke zjednodušení formulací některých tvrzení a předpokladů. Rozmanitost typů zpožděných diferenciálních rovnic a nutných specifických přístupů při jejich numerickém řešení přispívá k velkému rozvoji této teorie v posledních desetiletích.

V této práci se podařilo nashromáždit potřebnou teorii k numerickému řešení nejjednoduššího typu ZDR a na základě těchto poznatků by čtenář měl mít k dispozici všechny informace pro sestavení vlastní numerické metody pro řešení ZDR.

## Reference

- [1] Herbert Arndt. The influence of interpolation on the global error in retarded differential equations. In *Differential-Difference Equations/Differential-Differenzengleichungen*, pages 9–17. Springer, 1983.
- [2] Alfredo Bellen and Marino Zennaro. Stability properties of interpolants for runge-kutta methods. *SIAM journal on numerical analysis*, 25(2):411–432, 1988.
- [3] Alfredo Bellen and Marino Zennaro. *Numerical methods for delay differential equations*. Oxford university press, 2013.
- [4] SP Corwin, D Sarafyan, and S Thompson. Dk1ag6: A code based on continuously imbedded sixth-order runge-kutta methods for the solution of state-dependent functional differential equations. *Applied Numerical Mathematics*, 24(2-3):319–330, 1997.
- [5] John R Dormand and Peter J Prince. A family of embedded runge-kutta formulae. *Journal of computational and applied mathematics*, 6(1):19–26, 1980.
- [6] Rodney D Driver. Existence theory for a delay-differential system. *Contributions to Differential Equations*, 1(3):317–366, 1963.
- [7] L.E. El’sgol’ts and S.B. Norkin. *Introduction to the theory and application of differential equations with deviating arguments*, volume 105. Academic Press, 1973.
- [8] Wayne H Enright. Analysis of error control strategies for continuous runge-kutta methods. *SIAM Journal on Numerical Analysis*, 26(3):588–599, 1989.
- [9] Wayne H Enright. A new error-control for initial value solvers. *Applied Mathematics and Computation*, 31:288–301, 1989.
- [10] Wayne H Enright, Desmond J Higham, Brynjulf Owren, and Sharp. A survey of the explicit runge-kutta method. *Technical Report 291/94*, 1994.
- [11] WH Enright and H Hayashi. Convergence analysis of the solution of retarded and neutral delay differential equations by continuous numerical methods. *SIAM journal on numerical analysis*, 35(2):572–585, 1998.
- [12] Thomas Erneux. *Applied delay differential equations*. Surveys and tutorials in the applied mathematical sciences ; v. 3. Springer, New York, 2009.
- [13] Erwin Fehlberg. *Low-order classical Runge-Kutta formulas with stepsize control and their application to some heat transfer problems*, volume 315. National aeronautics and space administration, 1969.
- [14] S Filippi and U Buchacker. Stepsize control for delay differential equations using a pair of formulae. *Journal of computational and applied mathematics*, 26(3):339–343, 1989.
- [15] Nicola Guglielmi. Delay dependent stability regions of  $\theta$ -methods for delay differential equations. *IMA Journal of Numerical analysis*, 18(3):399–418, 1998.
- [16] Nicola Guglielmi and Ernst Hairer. Geometric proofs of numerical stability for delay equations. *IMA journal of numerical analysis*, 21(1):439–450, 2001.

- [17] Nicola Guglielmi and Ernst Hairer. Implementing radau iia methods for stiff delay differential equations. *Computing*, 67(1):1–12, 2001.
- [18] Jack K Hale. *Theory of functional differential equations*. Springer New York, 1977.
- [19] JK Hale. Homoclinic orbits and chaos in delay equations. In *Proceedings of the Ninth Dundee Conference on Ordinary and Partial Differential Equations*. New York: Wiley, 1986.
- [20] Desmond J Higham and Ioannis Th Famelis. Equilibrium states of adaptive algorithms for delay differential equations. *Journal of computational and applied mathematics*, 58(2):151–169, 1995.
- [21] Desmond J Higman. Error control for initial value problems with discontinuities and delays. *Applied numerical mathematics*, 12(4):315–330, 1993.
- [22] Mary Kathleen Horn. Fourth-and fifth-order, scaled rungs–kutta algorithms for treating dense output. *SIAM Journal on Numerical Analysis*, 20(3):558–568, 1983.
- [23] Kuang Jiaoxun. *Stability of numerical methods for delay differential equations*. Science Press, Beijing, 2005.
- [24] Vladimir Borisovich Kolmanovskii and Valerij Romanovič Nosov. *Stability of functional differential equations*, volume 180. Elsevier, 1986.
- [25] Stefano Maset. Stability of runge-kutta methods for linear delay differential equations. *Numerische Mathematik*, 87(2):355–371, 2000.
- [26] N.D.Hayes. Roots of the transcendental equation associated with a certain difference-differential equation. *Journal of the London Mathematical Society*, 1(3):226–232, 1950.
- [27] Ju Neimark. D-subdivisions and spaces of quasi-polynomials. *Prikladnaya Matematika i Mekhanika*, 13(5):349–380, 1949.
- [28] Kenneth W Neves and Alan Feldstein. Characterization of jump discontinuities for state dependent delay differential equations. *J. Math. Anal. Appl*, 56(3):689–707, 1976.
- [29] Jesper Ooppelstrup. The rkf4b method for delay—differential equations. In *Numerical treatment of differential equations*, pages 133–146. Springer, 1978.
- [30] D Sarafyan. Improved sixth-order runge-kutta formulas and approximate continuous solution of ordinary differential equations. *Journal of Mathematical Analysis and Applications*, 40(2):436–445, 1972.
- [31] D Sarafyan. New algorithms for the continuous approximate solutions of ordinary differential equations and the upgrading of the order of the processes. *Computers and Mathematics with Applications*, 20(1):77–100, 1990.
- [32] Diran Sarafyan. Continuous approximate solution of ordinary differential equations and their systems. *Computers and mathematics with applications*, 10(2):139–159, 1984.

- [33] Hans J Stetter. Considerations concerning a theory for ode-solvers. In *Numerical Treatment of Differential Equations*, pages 188–200. Springer, 1978.
- [34] Hans J Stetter. Tolerance proportionality in ode-codes. In *Proc. Second Conf. on Numerical Treatment of Ordinary Differential Equations*, pages 109–123. Humboldt University Berlin, 1980.
- [35] Marino Zennaro. One-step collocation: uniform superconvergence, predictor-corrector method, local error estimate. *SIAM journal on numerical analysis*, 22(6):1135–1152, 1985.
- [36] Marino Zennaro. Natural continuous extensions of runge-kutta methods. *Mathematics of Computation*, 46(173):119–133, 1986.



## Seznam zkratek

V práci byly uvedené tyto zkratky:

- ODR - obyčejné diferenciální rovnice;
- ZDR - zpožděné diferenciální rovnice;
- PSP - přirozené spojitě prodloužení;
- RK - Rungeovy-Kuttovy metody;
- SRK - spojitě Rungeovy-Kuttovy metody;



# Seznam příloh

V Příloze lze nalézt tyto numerické metody:

- Bellmanova metoda kroků, implementovaná pomocí dopředné Eulerovy metody s konstantním krokem, název souboru: “Bellman\_forward\_euler”;
- Feldsteinova metoda s první a druhou formulí, název souboru: “Feldstein”;
- Metoda kroků, implementovaná pomocí Eulerovy implicitní metody a lichobežníkové metody s konstantním krokem, název souboru: “metoda\_kroku\_implicitni\_metody”;
- Radau IIA Ehleova metoda s konstantním krokem, název souboru: “Radau\_IIA\_Ehle\_metoda\_RK\_pro\_ZDR”;
- Metoda DKLAG6 s konstantním krokem, název souboru: “DKLAG6”;
- RK metoda s variabilním krokem, název souboru: “variabilni\_krok\_RK\_pro\_ZDR”;