



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA STROJNÍHO INŽENÝRSTVÍ

FACULTY OF MECHANICAL ENGINEERING

ÚSTAV MATEMATIKY

INSTITUTE OF MATHEMATICS

**NUMERICKÉ ŘEŠENÍ ALGEBRAICKO-DIFERENCIÁLNÍCH
ROVNIC S INDEXEM 2**

NUMERICAL SOLUTION OF INDEX-2 DIFFERENTIAL-ALGEBRAIC EQUATIONS

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Tereza Kroulíková

VEDOUCÍ PRÁCE

SUPERVISOR

Mgr. Jitka Zatočilová, Ph.D.

BRNO 2016

Zadání bakalářské práce

Ústav:	Ústav matematiky
Studentka:	Tereza Kroulíková
Studijní program:	Aplikované vědy v inženýrství
Studijní obor:	Matematické inženýrství
Vedoucí práce:	Mgr. Jitka Zatočilová, Ph.D.
Akademický rok:	2015/16

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:

Numerické řešení algebraicko-diferenciálních rovnic s indexem 2

Stručná charakteristika problematiky úkolu:

Algebraicko-diferenciální rovnice se odlišují od obyčejných diferenciálních rovnic především tím, že systém algebraicko-diferenciálních rovnic neobsahuje derivace podle všech svých proměnných. Přešněji systém rovnic $F(t,x,x')=0$ nazveme algebraicko-diferenciálním pokud je matice dF/dx' singulární.

V průběhu zkoumání soustav algebraicko-diferenciálních rovnic došlo k vytvoření důležitého pojmu, totiž tzv. indexu DAE. Tato práce se bude proto zabývat numerickým řešením algebraicko-diferenciálních rovnic s indexem 2.

Cíle bakalářské práce:

1. Rozdělení algebraicko-diferenciálních rovnic podle indexu
2. Numerické řešení algebraicko-diferenciálních rovnic s indexem 2

Seznam literatury:

Uri M. Ascher, Linda R. Petzold - Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations

Peter Kunkel, Volker Mehrmann - Differential-Algebraic Equations: Analysis and Numerical Solution

Ernst Hairer, Gerhard Wanner - Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems

Termín odevzdání bakalářské práce je stanoven časovým plánem akademického roku 2015/16

V Brně, dne

L. S.

prof. RNDr. Josef Šlapal, CSc.
ředitel ústavu

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

Abstrakt

Práce se zabývá numerickým řešením algebraicko-diferenciálních rovnic. Tyto rovnice jsou nejprve popsány teoreticky a jsou ukázány jejich základní vlastnosti. Pozornost je věnována zejména indexu, jsou popsány nejpoužívanější indexy. Numerické řešení se zaměřuje na Hessenbergovy tvary algebraicko-diferenciálních rovnic indexu dva. Jsou zde odvozeny implicitní Runge-Kuttovy metody a metody zpětného derivování, které se používají pro řešení algebraicko-diferenciálních rovnic indexu 2.

Summary

This bachelor's thesis deals with numerical solution of differential-algebraic equations. At first these equations are described theoretically and their basic properties are presented. Main attention is paid to index and the most used indexes are described in details. Then the thesis concentrates on numerical solution of Hessenberg forms index-2 differential-algebraic equations. Implicit Runge-Kutta methods and backward differentiation formulas are derived. Those are used for solution of index-2 differential-algebraic equations.

Klíčová slova

Algebraicko-diferenciální rovnice, index, implicitní Runge-Kuttovy metody, metody zpětného derivování

Keywords

Differential-algebraic equations, index, implicit Runge-Kutta methods, backward differentiation formulas

KROULÍKOVÁ, T. *Numerické řešení algebraicko-diferenciálních rovnic s indexem 2*. Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2016. 43 s. Vedoucí Mgr. Jitka Zatočilová, Ph.D.

Prohlašuji, že jsem bakalářskou práci *Numerické řešení algebraicko-diferenciálních rovnic s indexem 2* vypracovala samostatně pod vedením Mgr. Jitky Zatočilové, Ph.D. s použitím materiálů uvedených v seznamu použité literatury.

Tereza Kroulíková

Děkuji svému školiteli Mgr. Jitce Zatočilové, Ph.D. za její užitečné rady a připomínky a za trpělivost při vedení mé bakalářské práce.

Tereza Kroulíková

Obsah

1	Úvod	13
2	Klasifikace DAE	15
2.1	Lineární DAE s konstantními koeficienty	15
2.2	Lineární DAE s nekonstantními koeficienty	17
2.3	Nelineární DAE	21
3	Index	23
4	Numerické řešení pro DAE	28
4.1	Implicitní Runge-Kuttovy metody	30
4.2	Metody zpětného derivování	35
5	Numerické řešení	38
6	Závěr	41

1. Úvod

Algebraicko-diferenciální rovnice se liší od obyčejných diferenciálních rovnic především tím, že systém algebraicko-diferenciálních rovnic neobsahuje derivace podle všech svých proměnných. Dříve se označovali jako singular, implicit, differential-algebraic descriptor, generalized state space, noncanonic, degenerate, semistate, reduced order model aj., podle toho, ve kterém oboru se problém formuloval. Nejvýpovědnějším názvem je constrained differential equations, tj. diferenciální rovnice s algebraickými omezeními. Tyto rovnice byly nakonec shrnuty pod pojem algebraicko-diferenciální rovnice, zkráceně DAE. Staly se velmi používaným nástrojem pro modelování a simulace omezených dynamických systémů jako jsou simulace elektrických obvodů, mechanických systémů, chemických reakcí, fluidní dynamiky a dalších, viz [4].

DAE byli dlouhou dobu řešeny numerickými metodami určenými pro obyčejné diferenciální rovnice (ODR). Ovšem ukazovalo se, že tyto metody nejsou vždy účinné, někdy dokonce selhávali u zdánlivě jednoduchých DAE. První teoretické práce se objevují v letech 1969 (Silverman) a 1977 (Luenberger). Průlomem byl článek *The simultaneous numerical solution of differential-algebraic equations* od C.W.Geara, který jako první aplikoval na některé třídy DAE metody zpětných derivací (BDF). Tento a ještě několik dalších článků byli popudem nejen k teoretickému studiu DAE, ale i k hledání nových metod řešení.

Bylo publikováno několik prací, které analyzují problém systémů DAE a vymezují třídy úloh, kde lze použít některé metody určené pro ODR. Klasifikace probíhala podle různých strukturálních vlastností. Došlo k vytvoření velmi důležitého pojmu, tzv. *indexu* DAE. Index DAE charakterizuje nakolik je systém obtížné řešit. Ukázalo se, že DAE s nižším indexem jsou snáze řešitelné než DAE s indexem 3 a vyšším [13].

Tato práce se věnuje numerickému řešení DAE s indexem 2. Nejprve budou DAE popsány teoreticky. Ukážeme některé jejich vlastnosti, zaměříme se zejména na index DAE a uvedeme nejčastěji používané typy. V druhé části se bude popsáno numerické řešení semi-explicitních systémů, konkrétně v Hessenbergově tvaru indexu 2. Tyto systémy lze řešit některými metodami pro řešení obyčejných diferenciálních rovnic. Budou odvozeny implicitní Runge-Kuttovy metody a metody zpětného derivování, které lze úspěšně použít právě pro řešení DAE.

2. Klasifikace DAE

Definice 2.1. Necht $(y, t) \in \mathcal{D} \times \mathcal{I}$, kde $\mathcal{D} \subseteq \mathbb{R}^n$ a $\mathcal{I} \subseteq \mathbb{R}$. Pak implicitní obyčejnou diferenciální rovnici

$$F(y', y, t) = 0 \quad (2.1)$$

nazýváme *algebraicko-diferenciální rovnici* (DAE), jestliže $F \in \mathcal{C}(\mathbb{R}^n \times \mathcal{D} \times \mathcal{I}, \mathbb{R}^n)$, existují parciální derivace $\frac{\partial}{\partial y} F(y', y, t)$, $\frac{\partial}{\partial y'} F(y', y, t)$ a Jakobián $\frac{\partial F}{\partial y'}$ je singulární s konstantní hodnotí pro všechna $(y', y, t) \in (\mathbb{R}^n \times \mathcal{D} \times \mathcal{I})$.

Implicitní tvar (2.1) lze velmi často přepsat následujícího názornějšího tvaru

$$G(y', y, z, t) = 0 \quad (2.2)$$

$$H(y, z, t) = 0, \quad (2.3)$$

kde (2.2) je soustava obyčejných diferenciálních rovnic a (2.3) jsou algebraická omezení. Protože se v soustavě nevyskytuje derivace podle proměnné z , nazýváme ji *algebraická proměnná*.

Stejně jako pro obyčejné diferenciální rovnice i u DAE se uvažuje počáteční úloha, kde musí být splněna nejen rovnice (2.1), ale i počáteční podmínka

$$y(t_0) = y_0, \quad t_0 \in \mathcal{I}, y_0 \in \mathbb{R}^n, \quad (2.4)$$

nebo okrajová úloha, která navíc splňuje okrajové podmínky

$$h(y(a), y(b)) = 0, \quad h \in \mathcal{D} \times \mathcal{D} \rightarrow \mathbb{R}^n, \mathcal{I} = \langle a, b \rangle.$$

Tyto podmínky musí být navíc v případě DAE konzistentní, tedy nesmí být v rozporu s algebraickými omezeními systému (2.1), viz [9].

My se budeme v tomto textu zabývat pouze počáteční úlohou pro DAE. Nyní uvedeme některé důležité typy DAE a jejich základní vlastnosti.

2.1. Lineární DAE s konstantními koeficienty

Nejjednodušším a nejlépe prostudovaným případem DAE jsou *lineární DAE s konstantními koeficienty*. Lze je řešit čistě algebraickými technikami. Zapisujeme je ve tvaru

$$Ay'(t) + By(t) = f(t), \quad (2.5)$$

kde A, B jsou čtvercové matice reálných nebo komplexních čísel a t je reálná proměnná. Většinu vlastností lineárních DAE s konstantními koeficienty lze odvodit právě z vlastností matic A, B .

Definice 2.2. Mějme matice A, B a komplexní parametr λ . Jestliže polynom $\det(\lambda A + B)$ není identicky rovný nule, pak $\lambda A + B$ nazýváme *regulární svazek*. V opačném případě mluvíme o svazku *singulárním*.

Věta 2.1. [4] *Algebraicko-diferenciální rovnice s konstantními koeficienty je řešitelná právě tehdy, když $\lambda A + B$ je regulární svazek.*

2.1. LINEÁRNÍ DAE S KONSTANTNÍMI KOEFICIENTY

Věta 2.2. Předpokládejme, že $\lambda A + B$ je regulární svazek, pak existují neregulární matice P, Q takové, že

$$PAQ = \begin{pmatrix} I & 0 \\ 0 & N \end{pmatrix}, PBQ = \begin{pmatrix} C & 0 \\ 0 & I \end{pmatrix} \quad (2.6)$$

kde N je nilpotentní matice řádu k , která je blokově diagonální a je tvořena Jordanovými bloky

$$\begin{pmatrix} 0 & 1 & 0 & \cdot & 0 \\ 0 & 0 & 1 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

a I je jednotková matice. Jestliže $N = 0$ pak definujeme $k = 0$. Ve speciálním případě, kdy A je regulární, vezmeme $PAQ = I$, $PBQ = C$ a definujeme $k = 0$. Jestliže $\det(\lambda A + B)$ je identicky konstantní, potom (2.6) zjednodušíme na $PAQ = N$ a $PBQ = I$.

Důkaz. Viz [10]. □

Přičemž řekneme, že matice N je nilpotentní matice řádu k jestliže $N^k = 0$ a $N^{k-1} \neq 0$. Řád nilpotence je také index svazku $\lambda A + B$ a také index algebraicko-diferenciální rovnice tzv. Kroneckerův index. Index je velmi důležitou vlastností DAE. Existuje několik typů indexů, Kroneckerův index je jedním z nich. O některých dalších typech indexů se zmíníme v kapitole 3.

Jestliže položíme $y = Qx$ a vynásobíme soustavu (2.5) zleva maticí P , kde P, Q jsou regulární matice, dostáváme

$$PAQx' + PBQx = Pf, \quad (2.7)$$

což je systém ekvivalentní systému (2.5). Vzhledem k (2.6) pak můžeme (2.7) zapsat ve tvaru

$$x_1' + Cx_1 = f_1 \quad (2.8)$$

$$Nx_2' + x_2 = f_2. \quad (2.9)$$

Soustava rovnic (2.8) je řešitelná pro libovolné počáteční podmínky. Na druhé straně soustava (2.9) je jednoznačně řešitelná pouze tehdy, pokud jsou počáteční podmínky konzistentní. V tomto případě lze řešení (2.9) zapsat ve tvaru

$$x_2 = (ND + I)^{-1} f_2 = \sum_{i=0}^{k-1} (-1)^i N^i f_2^{(i)}$$

kde $D = d/dt$, $f_2^{(i)} = d^i f_2 / dt^i$ a k je index soustavy.

Tento výpočet ilustruje některé vlastnosti řešení obecné soustavy algebraicko-diferenciálních rovnic:

1. řešení (2.5) může zahrnovat derivace řádu $k - 1$ funkce pravé strany f pokud je DAE vyššího indexu

2. ne všechny počáteční podmínky (2.5) připouští hladké řešení pro $k \geq 1$. Ty které řešení připouští se nazývají konzistentní počáteční podmínky.
3. DAE vyššího indexu mohou mít další skrytá algebraická omezení

V případě, že svazek $\lambda A + B$ je singulární, tedy charakteristický polynom je identicky rovný nule nebo matice A, B nejsou čtvercové, pak příslušná počáteční úloha má více než jedno řešení nebo její řešení vůbec neexistuje. Více lze nalézt v např. [9].

2.2. Lineární DAE s nekonzstantními koeficienty

Přirozeným zobecněním lineárních DAE s konstantními koeficienty jsou *lineární DAE s nekonzstantními koeficienty*, které jsou důležitým východiskem pro pochopení obecných nelineárních DAE. Zapisujeme je ve tvaru

$$A(t)y'(t) + B(t)y(t) = f(t) \quad (2.10)$$

kde $A(t), B(t)$ jsou čtvercové matice a $A(t)$ je navíc singulární pro všechna $t \in \mathcal{I}$.

Pro lineární DAE s konstantními koeficienty je regularita svazku $\lambda A + B$ a řešitelnost rovnice (2.5) ekvivalentní. Přirozeně proto chceme regularitu svazku zavést i pro lineární DAE s nekonzstantními koeficienty

Definice 2.3. Řekneme, že DAE (2.10) je *regulární*, jestliže svazek $\lambda A(t) + B(t)$ je regulární pro každé $t \in \mathcal{I}$.

Definice 2.4. Řekneme, že systém (2.10), kde A, B jsou matice typu $m \times m$, je řešitelný na intervalu \mathcal{I} , jestliže pro každou m -krát diferencovatelnou funkci f , existuje alespoň jedno spojitě diferencovatelné řešení (2.10). Kromě toho jsou řešení definována všude na \mathcal{I} a jsou jednoznačně určena pro každé $t \in \mathcal{I}$.

Bohužel, v případě lineárních DAE s nekonzstantními koeficienty již ekvivalence mezi řešitelností DAE a regularitou svazku $\lambda A + B$ neplatí. Existují regulární systémy, které nejsou řešitelné a naopak systémy singulární, které mají řešení. Jak ukazují následující dva příklady.

Příklad 2.1. Mějme soustavu (2.10), kde

$$A(t) = \begin{pmatrix} 0 & 0 \\ 1 & -t \end{pmatrix}, \quad B(t) = \begin{pmatrix} 1 & -t \\ 0 & 0 \end{pmatrix}, \quad f(t) = \begin{pmatrix} f_1(t) \\ f_2(t) \end{pmatrix}$$

Pak

$$\det(\lambda A(t) + B(t)) = -\lambda t + \lambda t = 0.$$

Svazek $\lambda A(t) + B(t)$ je tedy singulární pro všechna $t \in \mathcal{I}$. Položíme-li $y = (y_1, y_2)^T$, tak můžeme danou soustavu přepsat na následující tvar

$$\begin{aligned} 0 &= -y_1(t) + ty_2(t) + f_1(t) \\ y_1'(t) - ty_2'(t) &= f_2(t) \end{aligned}$$

2.2. LINEÁRNÍ DAE S NEKONSTANTNÍMI KOEFICIENTY

Z první rovnice dostaneme $y_1(t) = ty_2(t) + f_1(t)$. Jejím zderivováním a dosazením do druhé rovnice získáme jednoznačné řešení zadaného systému

$$\begin{aligned}y_1(t) &= tf_2(t) - f_1'(t) + f_1(t) \\y_2(t) &= f_2(t) - f_1'(t).\end{aligned}$$

Příklad 2.2. Nyní necht' A, B a f v (2.5) jsou dány následovně

$$A(t) = \begin{pmatrix} -t & t^2 \\ -1 & t \end{pmatrix}, B(t) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, f(t) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

potom $\det(\lambda A(t) + B(t)) = (1 - \lambda t)(1 + \lambda t) + t^2 \lambda^2 = 1$. Tedy svazek $\lambda A + B$ je regulární pro všechna $t \in \mathcal{I}$. Po krátkém výpočtu vidíme, že

$$y(t) = c(t) \begin{pmatrix} t \\ 1 \end{pmatrix}$$

splňuje rovnici (2.10) pro libovolnou skalární funkci $c(t) \in \mathcal{C}^1(\mathbb{R}, \mathbb{C})$, s $c(t_0) = 0$. Tato DAE je sice regulární, ovšem není jednoznačně řešitelná, protože počáteční podmínky neurčují řešení jednoznačně.

Protože index DAE hraje důležitou úlohu při výběru vhodné numerické metody, tak i zde tento pojem zavedeme pomocí regulárního svazku matic $A(t), B(t)$. Protože ale matice $A(t)$ a $B(t)$ v (2.10) nejsou konstantní, bude mít pouze lokální charakter.

Definice 2.5. Jestliže (2.10) je regulární algebraicko-diferenciální rovnice potom *lokálním indexem* v t , značíme $\nu_l(t)$, je index svazku $\lambda A(t) + B(t)$.

Stejně jako v případě lineárních DAE s konstantními koeficienty můžeme (2.10) převést na ekvivalentní systém a to tak, že položíme $y = Q(t)x$ a vynásobíme (2.5) zprava maticí $P(t)$. Matice $P(t), Q(t)$ jsou regulární pro $t \in \mathcal{I}$ a alespoň tak hladké jako koeficienty DAE. Získáváme tím systém

$$PAQx' + (PAQ' + PBQ)x = Pf, \quad (2.11)$$

který je ekvivalentní systému (2.5).

Jaký je vztah tohoto lokálního indexu ν_l a celkového indexu ν , který nás zajímá především, ukazují následující věty.

Věta 2.3. *Předpokládáme, že (2.10) je řešitelná DAE na intervalu \mathcal{I} . Potom následující tvrzení jsou ekvivalentní*

1. (2.10) je implicitní ODR
2. $\nu = 0$
3. $\nu_l \equiv 0$
4. $\nu_l = 0$ pro nějaké $t_0 \in \mathcal{I}$

kde ν je index a ν_l je lokální index (2.10).

Věta 2.4. Předpokládáme, že (2.10) je řešitelná. Potom $\nu = 1$ právě tehdy když $\nu_l = 1$

Dvě předchozí věty ukazují, že ν a ν_l jsou stejné pro index nula a index jedna pro řešitelné DAE. Toto ovšem neplatí pro systémy vyšších indexů ($\nu \geq 2$).

Protože při numerickém řešení je třeba často pracovat se systémem ekvivalentním k systému (2.10), uvedeme ještě následující větu.

Věta 2.5. Předpokládáme, že (2.10) je řešitelná DAE na intervalu \mathcal{I} . Pokud transformujeme (2.10) na ekvivalentní systém (2.11), tak zůstane zachováno

1. $\nu_l \equiv 0$
2. $\nu_l \equiv 1$
3. $\nu_l(t) \geq 2$ pro všechna $t \in \mathcal{I}$.

Pokud je $\nu_l \geq 2$ na \mathcal{I} potom je možné najít analyticky ekvivalentní systém takový, že $\nu_l = 2$ na otevřené husté podmnožině \mathcal{I} [4]. Tyto výsledky ukazují, že můžeme uvažovat tři přirozené třídy řešitelných DAE - implicitní ODR, DAE indexu 1 a DAE indexu 2.

Systém (2.10) se velmi často vyskytuje ve speciálním tvaru, který je označován jako semi-explicitní tvar

$$\begin{aligned} y' + B_{11}(t)y + B_{12}(t)z &= f_1 \\ B_{21}(t)y + B_{22}(t)z &= f_2. \end{aligned} \quad (2.12)$$

Mnoho semi-explicitních DAE, které se vyskytují v aplikacích, mají navíc určitou strukturu.

Definice 2.6. Řekneme, že lineární DAE s nekonstantními koeficienty (2.10) je v Hessenbergově tvaru velikosti r , jestliže (2.10) může být zapsána ve tvaru

$$\begin{pmatrix} I & 0 & \cdot & \cdot & 0 \\ 0 & I & \cdot & \cdot & \cdot \\ \cdot & \cdot & I & \cdot & \cdot \\ \cdot & \cdot & \cdot & I & \cdot \\ 0 & \cdot & \cdot & \cdot & 0 \end{pmatrix} \begin{pmatrix} y'_1 \\ \cdot \\ \cdot \\ \cdot \\ y'_r \end{pmatrix} + \begin{pmatrix} B_{11}(t) & * & * & B_{1,r-1}(t) & B_{1r}(t) \\ B_{21}(t) & * & * & B_{2,r-1}(t) & \cdot \\ 0 & * & * & * & \cdot \\ \cdot & \cdot & * & * & \cdot \\ 0 & \cdot & 0 & B_{r,r-1}(t) & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_r \end{pmatrix} = \begin{pmatrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_r \end{pmatrix},$$

kde y_i jsou vektory, B_{ij} matice a součin matic $B_{r,r-1}B_{r-1,r-2}\dots B_{1,r}$ je regulární.

Poznamenejme, že obecně matice $B_{i+1,i}$, $i = 1, \dots, 1 - r$ nejsou čtvercové. Pouze jejich součin musí být čtvercová a regulární matice.

Jak ukazuje následující věta velikost Hessenbergova tvaru je také úzce spjata s indexem soustavy.

Věta 2.6. [4] DAE v Hessenbergově tvaru velikosti r je řešitelná a má index a lokální index r .

2.2. LINEÁRNÍ DAE S NEKONSTANTNÍMI KOEFICIENTY

Nejčastěji se v praxi vyskytuje Hessenbergův tvar velikosti 2 nebo 3 a proto je zde uvedeme. Hessenbergův tvar velikosti (indexu) dva je tvaru

$$\begin{aligned} y' + B_{11}(t)y + B_{12}(t)z &= f_1 \\ B_{21}(t)y &= f_2, \end{aligned} \quad (2.13)$$

kde součin matic $B_{21}B_{12}$ je regulární. Hessenbergův tvar velikosti tři je pak tvaru

$$\begin{aligned} y'_1 + B_{11}(t)y_1 + B_{12}(t)y_2 + B_{13}(t)y_3 &= f_1 \\ y'_1 + B_{21}(t)y_1 + B_{22}(t)y_2 &= f_2 \\ B_{32}(t)y_2 &= f_3, \end{aligned} \quad (2.14)$$

kde je regulární součin matic $B_{32}B_{21}B_{13}$.

Nyní uvedeme příklad lineární DAE s nekonstantními koeficienty v Hessenbergově tvaru. Tento příklad bude numericky řešen v kapitole 5.

Příklad 2.3. Příkladem lineární DAE s nekonstantními koeficienty v Hessenbergově tvaru s indexem 2 je soustava:

$$\begin{aligned} y'_1 &= \left(\alpha - \frac{1}{2-t}\right)y_1 + (2-t)\alpha z + \frac{3-t}{2-t}e^t \\ y'_2 &= \frac{1-\alpha}{t-2}y_1 - y_2 + (\alpha-1)z + 2e^t \\ 0 &= (t+2)y_1 + (t^2-4)y_2 - (t^2+t-2)e^t \end{aligned} \quad (2.15)$$

kde α je parametr. Přepíšeme-li (2.15) do Hessenbergova tvaru (2.13):

$$\begin{aligned} y'_1 \quad -\left(\alpha - \frac{1}{2-t}\right)y_1 \quad & \quad \quad \quad -(2-t)\alpha z &= \frac{3-t}{2-t}e^t \\ y'_2 \quad \quad -\frac{1-\alpha}{t-2}y_1 \quad \quad -y_2 \quad -(\alpha-1)z &= 2e^t \\ (t+2)y_1 \quad + (t^2-4)y_2 & \quad \quad \quad &= (t^2+t-2)e^t \end{aligned}$$

pak vidíme, že matice B_{11} , B_{12} a B_{21} mají tvar

$$B_{11} = \begin{pmatrix} \frac{1}{2-t} - \alpha & 0 \\ \frac{\alpha-1}{t-2} & 1 \end{pmatrix}, \quad B_{12} = \begin{pmatrix} (t-2)\alpha \\ 1-\alpha \end{pmatrix}, \quad B_{21} = \begin{pmatrix} t+2, & t^2-4 \end{pmatrix}$$

a součin $B_{21}B_{12} = (t+2)(t-2)\alpha + (t^2-4)(1-\alpha) = (t^2-4)$ je regulární pro $t \neq \pm 2$.

Pro (2.15) s počátečními podmínkami $y_1(0) = y_2(0) = 1$ známe i přesné řešení

$$y_1 = y_2 = e^t, \quad z = -\frac{e^t}{2-t}.$$

Dalším tvarem lineárních DAE s nekonstantními koeficienty je kanonický tvar. DAE v kanonickém tvaru nebo v kombinaci kanonického tvaru a Hessenbergova jsou řešitelné. Dále se jím ovšem tato práce nebude zabývat, proto zde nebude uveden. Lze ho nalézt v [4, 9].

2.3. Nelineární DAE

Nejobecnější DAE jsou nelineární. Zapisujeme je ve tvaru

$$F(y', y, z, t) = 0. \quad (2.16)$$

Protože se budeme v dalších kapitolách, které se týkají numerického řešení, věnovat pouze DAE v semi-explicitním tvaru, tak se zde omezíme jen na takové typy nelineárních rovnic (2.16), které lze do semi-explicitního tvaru

$$\begin{aligned} y' &= f(y, z, t) \\ 0 &= g(y, z, t) \end{aligned} \quad (2.17)$$

přepsat. Obdobně jako v přechodí kapitole mohou mít i nelineární DAE v semi-explicitním tvaru specifickou strukturu.

Definice 2.7. Řekneme, že DAE (2.16) je v *Hessebergově tvaru velikosti r* , jestliže

$$\begin{aligned} y'_1 &= F_1(y_1, y_2, \dots, y_r, t) \\ y'_2 &= F_2(y_1, y_2, \dots, y_{r-1}, t) \\ &\vdots \\ y'_i &= F_i(y_1, y_2, \dots, y_{r-i}, t), \quad 3 \leq i \leq r-1 \\ &\vdots \\ 0 &= F_r(y_{r-1}, t), \end{aligned} \quad (2.18)$$

kde $\frac{\partial F_r}{\partial y_{r-1}} \frac{\partial F_{r-1}}{\partial y_{r-2}} \dots \frac{\partial F_2}{\partial x_1} \frac{\partial F_1}{\partial x_r}$ je regulární.

Hessenbergův tvar velikosti (indexu) 2 je tedy

$$\begin{aligned} y' &= f(y, z, t) \\ 0 &= g(y, t), \end{aligned} \quad (2.19)$$

kde $g(y, t)$ je m -rozměrný vektor, tedy g_y je matice $m \times n$. Na druhou stranu $f(y, z, t)$ je n -rozměrný vektor, tedy f_z je matice $n \times m$. Výsledek součinu $g_y f_z$ je tedy matice $m \times m$ a je regulární. A můžeme si povšimnout, že algebraické omezení nezávisí na algebraické proměnné.

Na závěr této kapitoly ještě uvedeme příklad nelineární DAE v Hessenbergově tvaru s indexem 2, který bude numericky řešen v kapitole 5.

Příklad 2.4. Příkladem DAE v Hessenbergově tvaru s indexem dva je soustava:

$$\begin{aligned} y'_1 &= y_1 y_2^2 z^2 \\ y'_2 &= y_1^2 y_2^2 - 3y_2^2 z \\ 0 &= y_1^2 y_2 - 1. \end{aligned} \quad (2.20)$$

Vypočteme Jakobiány g_y a f_z :

$$g_y = (2y_1 y_2, y_1^2), \quad f_z = (2y_1 y_2^2 z, -3y_2^2)^T$$

2.3. NELINEÁRNÍ DAE

a provedeme jejich součin

$$g_y f_z = 4y_1^2 y_2^3 z - 3y_1^2 y_2^2 = y_1^2 y_2^2 (4y_2 z - 3). \quad (2.21)$$

Pokud tedy budou splněny podmínky $y_1 \neq 0$ a zároveň $y_2 \neq 0$ a zároveň $y_2 z \neq \frac{3}{4}$ dostaneme regulární matici velikosti jedna.

Pro počáteční podmínky $y_1(0) = 1, y_2(0) = 1$ má tato úloha přesné řešení

$$y_1(t) = e^t, \quad y_2(t) = e^{-2t}, \quad z(t) = e^{2t}, \quad (2.22)$$

tedy je zřejmé, že výše uvedená podmínka regularity bude splněna.

3. Index

V průběhu výzkumu algebraicko-diferenciálních rovnic vznikl důležitý pojem a to tzv. *index* algebraicko-diferenciální rovnice. Index charakterizuje stupeň obtížnosti řešení soustavy. Obecně platí, že čím vyšší index soustavy, tím obtížnější je problém numericky řešit. Různými přístupy vznikly různé typy indexů.

Diferenciální index

Diferenciální index je nejznámějším indexem a může být považován za jakési měřítko nakolik je systém DAE vzdálený od ODR, které mají diferenciální index nula.

Definice 3.1. *Diferenciálním indexem* obecné soustavy DAE (2.1) nazveme minimální počet nutných derivování (podle t) částí nebo celého systému (2.1), abychom mohli určit y' jako spojitou funkci y a t .

Pro semi-explicitní systém (2.17) určíme diferenciální index následujícím způsobem: zderivujeme algebraická omezení podle t a dostaneme:

$$g_z(y, z, t)y' + g_y(y, z, t)z' = -g_t(y, z, t) \quad (3.1)$$

Jestliže je Jakobián g_y regulární, potom získáme implicitní ODR a soustava má index 1. Je-li g_z singulární potom algebraickými operacemi a změnami souřadnic můžeme převést opět na (3.1) semi-explicitní soustavu DAE (samozřejmě pro jiné proměnné y a z). Dále znovu derivujeme algebraické rovnice, a jestliže dostaneme implicitní ODR, potom původní systém má index 2 atd.

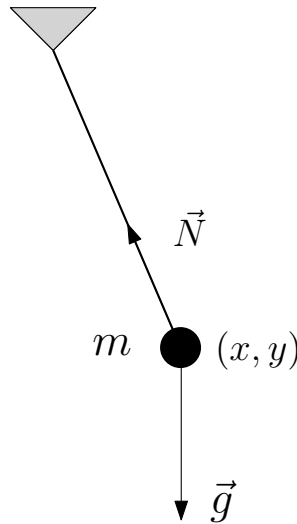
Nyní ukážeme jak určit diferenciální index na příkladu rovinného kyvadla.

Příklad 3.1. Fyzikální kyvadlo popisuje pohyb hmotného bodu o hmotnosti m v kartézských souřadnicích (x, y) , který je zavěšen na závěsu délky $l = \sqrt{x^2 + y^2}$. Na hmotný bod působí síla tíhová, která působí dolů velikostí $-mg$, a tahová, která působí ve směru tyče. Jestliže její velikost označíme N potom její složky jsou $(-Nx, -Ny)/\sqrt{x^2 + y^2}$. Dohromady tyto vztahy vytváří kompletní model kyvadla

$$\begin{aligned} m \frac{d^2x}{dt^2} &= -N \frac{x}{\sqrt{x^2 + y^2}} \\ m \frac{d^2y}{dt^2} &= -N \frac{y}{\sqrt{x^2 + y^2}} - mg \\ 0 &= l - \sqrt{x^2 + y^2} \end{aligned}$$

Tento systém druhého řádu může být přepsán jako systém prvního řádu

$$\begin{aligned} x' &= u \\ y' &= v \\ mu' &= -N \frac{x}{\sqrt{x^2 + y^2}} \\ mv' &= -N \frac{y}{\sqrt{x^2 + y^2}} - mg \\ 0 &= l - \sqrt{x^2 + y^2}. \end{aligned}$$



Obrázek 3.1: Model kyvadla.

Neznámé jsou souřadnice $x(t)$, $y(t)$, jejich rychlosti $u(t)$, $v(t)$ a síla vynaložená tyčí $N(t)$. Dohromady tedy máme pět rovnic o pěti neznámých. Přičemž čtyři z nich jsou diferenciální. Není zde žádná rovnice s N' , tedy N je algebraická proměnná.

Pro zjednodušení označíme $\lambda = N/(m\sqrt{x^2 + y^2})$, pak $u' = -\lambda x$ a $v' = \lambda y - g$. Také algebraickou omezující rovnici můžeme přepsat na $0 = l^2 - x^2 - y^2$. Nyní převedeme algebraicko-diferenciální rovnice na systém obyčejných diferenciálních rovnic tak, že budeme derivovat algebraickou rovnici, dokud nedostaneme vyjádření pro λ' . Derivování algebraické rovnice dává nejprve

$$0 = \frac{d}{dt}(l^2 - x^2 - y^2) = -2xu - 2yv \quad (3.2)$$

zderivováním (3.2) dostaneme

$$0 = \frac{d^2}{dt^2}(l^2 - x^2 - y^2) = -2(u^2 + v^2) + 2\lambda(x^2 + y^2) + 2yg \quad (3.3)$$

a teprve derivací (3.3) dostáváme

$$0 = \frac{d^3}{dt^3}(l^2 - x^2 - y^2) = -2\frac{d\lambda}{dt}(x^2 + y^2) + 6gv.$$

Nyní třikrát zderivovanou podmínku přidáme k diferenciálním rovnicím a získáváme soustavu obyčejných diferenciálních rovnic.

$$\begin{aligned} x' &= u \\ y' &= v \\ u' &= -\lambda x \\ v' &= -\lambda y - g \\ \lambda' &= -\frac{3gv}{l^2}. \end{aligned}$$

Takže původní systém DAE (3.2) má diferenciální index 3.

Rovnice (3.2) a (3.3) představují skrytá algebraická omezení, o kterých jsme už zmínili v závěru kapitoly 2.1, viz také [2].

Kroneckerův index

Kroneckerův index je nejstarším typem indexu. Byl definován především pro lineární DAE s konstantními koeficienty. Jak již bylo uvedeno tento index je definován jako řád nilpotence matice N v (2.6), proto se mu také někdy říká index nilpotence. Kroneckerův index koresponduje s diferenciálním indexem pro lineární DAE s konstantními koeficienty, jak ukazuje i následující příklad [4, 12].

Příklad 3.2. Mějme následující DAE s konstantními koeficienty

$$\begin{aligned}y_1' + y_3 &= g_1 \\y_2' + y_1 &= g_2 \\y_2 &= g_3\end{aligned}\tag{3.4}$$

postupným derivováním algebraické podmínky dostáváme vztah

$$y_2' = g_3'.$$

Dosazením do druhé rovnice soustavy (3.4) získáme

$$y_1 = g_2 - g_3'.$$

a tento vztah opět zderivujeme, čímž dostaneme

$$y_1' = g_2' - g_3''.$$

Tento vztah nyní dosadíme do první rovnice soustavy (3.4) a vyjádříme y_3

$$y_3 = g_1 - g_2' + g_3''$$

a teprve třetí derivací získáváme

$$y_3' = g_1' - g_2'' + g_3'''.$$

Tedy soustava (3.4) má diferenciální index 3.

Soustavu (3.4) ale můžeme zapsat i ve tvaru

$$Ay' + By = g,$$

kde

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ a } B = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

pro

$$P = I \text{ a } Q = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

je

$$PAQ = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = N,$$

protože $\det(\lambda A + B) = 1$, lze snadno ověřit krátkým výpočtem, a

$$PBQ = I.$$

Protože $N^2 \neq 0$ a $N^3 = 0$, tak Kroneckerův index je také 3.

Perturbační index

Perturbační index definoval E. Hairer, C. Lubich a M. Roche a je měřítkem citlivosti řešení vzhledem k perturbacím daného problému. Perturbace se mohou objevit vlivem zaokrouhlování a numerických aproximací. Proto je tento index v numerickém řešení důležitý a často užívaný.

Definice 3.2. DAE (2.1) má *perturbační index* k podél řešení $y^* \in \mathcal{C}^1(\mathcal{I}, \mathcal{D})$ na kompaktním intervalu $\mathcal{I} = \langle t_0, T \rangle$, jestliže k je nejmenší číslo takové, že pro všechny funkce $\bar{y} \in \mathcal{C}^1(\mathcal{I}, \mathcal{D})$ pro které

$$F(\bar{y}', \bar{y}, t) = \delta(t),$$

kde $\delta \in \mathcal{C}^{k-1}(\mathcal{I}, \mathbb{R})$ je defekt, je pro všechna $t \in \mathcal{I}$ splněna nerovnost

$$\|\bar{y} - y^*\|_\infty \leq c \left(\|\bar{y}(t_0) - y^*(t_0)\|_\infty + \sum_{j=0}^{k-1} \|\delta^{(j)}\|_\infty \right),$$

pokud výrazy na pravé straně jsou dostatečně malé, $c > 0$ je libovolné.

Jeho nedostatkem je, že jeho určení vyžaduje znalost přesného řešení. Dříve se předpokládalo, že mezi diferenciálním a perturbačním indexem platí následující vztah:

$$\text{diferenciální index} \leq \text{perturbační index} \leq \text{diferenciální index} + 1$$

ukázalo se, že tomu tak být nemusí, jak je ukázáno na následujícím příkladu

Příklad 3.3. Máme soustavu

$$y_m N \mathbf{y}' + \mathbf{y} = 0 \tag{3.5}$$

kde N je horní trojúhelníkový Jordanův blok velikosti $m \times m$. Poslední řádek N je nulový, takže máme $y_m = 0$ a diferenciální index 1. Na druhou stranu přidání perturbací způsobí, že y_m již nebude nulové. A to je důvod proč perturbační index (3.5) je m , viz [3, 8, 11].

Další typy indexů

Jako další méně používané indexy lze zmínit *strangeness index* zavedený P. Kunkelem a V. Mehrmannem, který je rozšířením diferenciálního indexu na pře- a pod-určené systémy a závisí na transformacích na kanonický tvar.

Geometrický index byl zaveden při studiu DAE jako diferenciálních rovnic na varietách a využívá se v metodách redukce.

Strukturální index je kombinatoricky orientovaný index, který byl nejprve definován pro lineární DAE s konstantními koeficienty. Je definován tak, že nahradíme-li v $\lambda A + B$ nenulové prvky nezávislými parametry p_j , potom jednoznačně určené celé číslo, které se rovná Kroneckerově indexu svazku $\lambda A(p) + B(p)$ pro všechna p se nazývá strukturálním indexem [11, 12].

4. Numerické řešení pro DAE

V této kapitole se budeme zabývat numerickým řešením DAE. Zaměříme se na počáteční problém. Numerickým řešením počáteční úlohy rozumíme výpočet přibližných hodnot y_n hledaného řešení $y(t)$ v bodech t_n , $n = 1, \dots, N$, které nazýváme uzly. Proto na intervalu $\langle a, b \rangle$ zavedeme dělení $a = t_0 < t_1 < \dots < t_N = b$ tak, aby dostatečně pokrylo tento interval. Vzdálenost dvou sousedních uzlů $h_n = t_{n+1} - t_n$ nazveme délkou kroku. My budeme uvažovat rovnoměrné dělení, tj. $h_n = h = (b - a)/N$, v tom případě $t_n = a + nh$.

K numerickému řešení DAE existují obecně dva přístupy. První je přímá diskretizace daného problému. Druhý zahrnuje reformulaci problému, např. redukci indexu, kombinovanou s diskretizací. My se v této práci budeme věnovat pouze přímým diskretizačním metodám.

Při určování jaké numerické metody by byly vhodné pro řešení DAE v semi-explicitním tvaru nám pomůže následující úvaha. Máme DAE v semi-explicitním tvaru (2.17), kde nahradíme rovnici $0 = g(y, z, t)$ následující rovnicí

$$\varepsilon z' = g(y, z, t)$$

a tím získáme ODR

$$\begin{aligned} y' &= f(y, z, t), & y(t_0) &= y_0 \\ \varepsilon z' &= g(y, z, t) \end{aligned} \quad (4.1)$$

závislou na parametru $0 \leq \varepsilon \ll 1$. Tato soustava je tuhá, proto explicitní metody nejsou vhodné, protože jsou nestabilní. Jestliže bychom chtěli použít explicitní metody, museli bychom volit velmi malý krok. Vhodnější jsou proto pro řešení tuhých systémů implicitní metody. Protože pro velmi malé ε je soustava (4.1) velmi blízká DAE, tak i k jejich řešení se také používají metody určené pro tuhé ODR, nejčastěji metody zpětného derivování a implicitní Runge-Kuttovy metody. Tyto metody byly odvozeny pro řešení ODR, na DAE se dají použít pouze za některých předpokladů. Především je nutné, aby byl problém matematicky dobře formulován, tj. aby soustava DAE byla řešitelná. Tedy metoda musí být zvolena tak, aby vzniklý nelineární systém rovnic byl řešitelný v každém kroku.

Jak jsme již zmínili nejvíce se používají metody zpětného derivování a implicitní Runge-Kuttovy metody. Nejjednodušším zástupcem obou skupin metod je implicitní Eulerova metoda. Použijeme-li ji pro (4.1) s krokem velikosti h , dostáváme

$$\begin{aligned} y_{n+1} &= y_n + hf(y_{n+1}, z_{n+1}), \\ \varepsilon z_{n+1} &= \varepsilon z_n + hg(y_{n+1}, z_{n+1}), \end{aligned}$$

Pro DAE indexu 1 dostaneme uspokojivý výsledek. Pro obecný systém DAE indexu 2 už situace není natolik jednoduchá. V následujícím příkladu ukážeme, že DAE indexu 2 lze úspěšně řešit výše uvedenými implicitními numerickými metodami především pokud je DAE dána v semi-explicitním tvaru.

Příklad 4.1. Máme následující lineární DAE s nekonstantními koeficienty indexu 2, která závisí na parametru η

$$\begin{pmatrix} 0 & 0 \\ 1 & \eta t \end{pmatrix} y' + \begin{pmatrix} 1 & \eta t \\ 0 & 1 + \eta \end{pmatrix} y = \begin{pmatrix} q(t) \\ 0 \end{pmatrix}. \quad (4.2)$$

Tato soustava má přesné řešení

$$\begin{aligned}y_1(t) &= q(t) + \eta tq'(t) \\y_2(t) &= -q'(t),\end{aligned}$$

které je definováno pro všechny hodnoty η . Problém je stabilní pro střední hodnoty η .

Nyní si transformací souřadnic rovnicí (4.2) převedeme na ekvivalentní semi-explicitní tvar tak, že položíme $y = Q(t)(u, v)^T$, kde $y = (y_1, y_2)^T$ a

$$Q(t) = \begin{pmatrix} 1 & -\eta t \\ 0 & 1 \end{pmatrix}.$$

Odtud pak

$$y = \begin{pmatrix} 1 & -\eta t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

a ekvivalentní systém k (4.2) tedy je

$$\begin{aligned}u' + v &= 0 \\u &= q(t).\end{aligned}\tag{4.3}$$

Implicitní Eulerova metoda nám pro (4.3) dává

$$\begin{aligned}u_{n+1} &= q(t_{n+1}) \\v_{n+1} &= -\frac{q(t_{n+1}) - u_n}{h}.\end{aligned}$$

Tedy zaručuje, že při konzistentní počáteční podmínce např. $u_0 = q(0)$ dostaneme

$$v_{n+1} = -q'(t_{n+1}) + \mathcal{O}(h)$$

což je vše co můžeme čekat od metod prvního řádu pro ODR.

Když ale použijeme implicitní Eulerovu metodu přímo na (4.2), obdržíme

$$\begin{pmatrix} 0 & 0 \\ 1 & \eta t_{n+1} \end{pmatrix} \frac{y_{n+1} - y_n}{h} + \begin{pmatrix} 1 & \eta t_{n+1} \\ 0 & 1 + \eta \end{pmatrix} y_{n+1} = \begin{pmatrix} -q(t_{n+1}) \\ 0 \end{pmatrix}.$$

Položíme-li nyní

$$\begin{pmatrix} u_{n+1} \\ v_{n+1} \end{pmatrix} = \begin{pmatrix} 1 & \eta t_{n+1} \\ 0 & 1 \end{pmatrix} y_{n+1}$$

získáváme následující diskterizaci

$$\begin{aligned}u_{n+1} &= q(t_{n+1}) \\(1 + \eta)v_{n+1} &= \frac{q(t_{n+1}) - q(t_n)}{h}.\end{aligned}$$

Vidíme, že zatímco u_{n+1} je určeno přesně, v_{n+1} není definováno pro $\eta = -1$ a pro $\eta \neq 0$ má chybu $\mathcal{O}(1)$.

4.1. IMPLICITNÍ RUNGE-KUTTOVY METODY

Převedením na semi-explicitní tvar se oddělí řešení části y na část diferenciální a algebraickou. Právě pro takto oddělené problémy pracuje implicitní Eulerova metoda velmi dobře. Obecně se přímá diskretizace neoddělených DAE vyššího indexu než jedna nedoporučuje.

My se dále budeme zabývat pouze semi-explicitními DAE, na které přímé diskretizační metody lze použít, a v následujících dvou kapitolách si dvě nejznámější popíšeme. Přímé diskretizační metody jsou omezeny na problémy indexu 1 a semi-explicitní systémy indexu 2. Naštěstí velká část DAE formulovaných v praktických problémech mají buď index 1 a nebo, pokud jde o problémy vyšších indexů, můžou být vyjádřeny jako kombinace Hessenbergových tvarů, viz [1].

4.1. Implicitní Runge-Kuttovy metody

Runge-Kuttovy metody patří mezi nejznámější jednokrokové metody pro řešení počátečních úloh ODR. Pro netuhé problémy se využívají spíše explicitní Runge-Kuttovy metody. V případě tuhých problémů a DAE sáhneme po implicitních Runge-Kuttových metodách (zkráceně IRK), jejichž předností je neomezená oblast absolutní stability. Nejprve si uvedme IRK metody pro ODR, odvodíme je a ukážeme základní vlastnosti. Poté teprve přejdeme k řešení DAE.

Implicitní Runge-Kuttovu metodu odvodíme z diferenciální rovnice

$$y'(t) = f(t, y(t)) \quad (4.4)$$

s počáteční podmínkou

$$y(t_0) = y_0 \quad (4.5)$$

tím, že (4.4) převedeme na integrální rovnici

$$\int_{t_n}^t y'(r) dr = \int_{t_n}^t f(r, y(r)) dr$$

a zintegrujeme pravou stranu na intervalu $\langle t_n, t \rangle$

$$y(t) = y(t_n) + \int_{t_n}^t f(r, y(r)) dr. \quad (4.6)$$

Nejprve ve vztahu (4.6) nahradíme $y(t_n)$ jeho přibližnou hodnotou y_n a poté nahradíme funkci $f(r, y(r))$ na intervalu $\langle t_n, t_{n+1} \rangle$ interpolačním polynomem stupně $< s$ procházejícím body $\{t_{n,i} \equiv t_n + \tau_i h : i = 1, \dots, s\}$, kde $0 \leq \tau_1 < \dots < \tau_s \leq 1$ jsou zvolené konstanty. Dostáváme tedy

$$y(t) \approx y_n + \int_{t_n}^t p(r) dr. \quad (4.7)$$

Použijeme-li Lagrangeův interpolační polynom, tak

$$p(r) = \sum_{j=1}^s f(t_{n,j}, y(t_{n,j})) l_j(r), \quad (4.8)$$

kde $l_j(r)$ jsou fundamentální Lagrangeovy polynomy definované předpisem

$$l_j(r) = \prod_{\substack{i=1 \\ i \neq j}}^s \left(\frac{r - r_i}{r_j - r_i} \right), \quad j = 1, \dots, s. \quad (4.9)$$

Dosazením (4.8) do (4.7) dostáváme

$$y(t) \approx y_n + \sum_{j=1}^s f(t_{n,j}, y(t_{n,j})) \int_{t_n}^t l_j(r) dr.$$

Nahradíme-li v předchozím vztahu $y(t_{n,j})$ jejich přibližnými hodnotami $y_{n,j}$ pro $j = 1, \dots, s$, tak dostáváme následující soustavu nelineárních rovnic

$$y_{n,i} = y_n + \sum_{j=1}^s f(t_{n,i}, y_{n,j}) \int_{t_n}^{t_{n,i}} l_j(r) dr, \quad i = 1, \dots, s. \quad (4.10)$$

Pokud $\tau_s = 1$, tak $y_{n,s} = y_{n+1}$ a tedy

$$y_{n+1} = y_n + \sum_{j=1}^s f(t_{n,j}, y_{n,j}) \int_{t_n}^{t_{n+1}} l_j(r) dr. \quad (4.11)$$

Ze vztahů (4.10) a (4.11) tedy plyne obecný tvar s-stupňové IRK metody

$$k_{n,i} = y_n + h \sum_{j=1}^s a_{i,j} f(t_n + c_j h, k_{n,j}), \quad i = 1, \dots, s, \quad (4.12)$$

$$y_{n+1} = y_n + h \sum_{j=1}^s b_j f(t_n + c_j h, k_{n,j}), \quad (4.13)$$

kde koeficienty $a_{i,j}, b_j, c_j$ lze zapsat do následující Butcherovy tabulky

c_1	$a_{1,1}$	$a_{1,2}$	\cdots	$a_{1,s-1}$	$a_{1,s}$	zkráceně		A
c_2	$a_{2,1}$	$a_{2,2}$	\cdots	$a_{2,s-1}$	$a_{2,s}$			
c_3	$a_{3,1}$	$a_{3,2}$	\cdots	$a_{3,s-1}$	$a_{3,s}$			
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots			
c_s	$a_{s,1}$	$a_{s,2}$	\cdots	$a_{s,s-1}$	$a_{s,s}$			
	b_1	b_2	\cdots	b_{s-1}	b_s			b^T

Integrály ve vztazích (4.10) a (4.11) lze jednoduše vyčíslit. Například pro $s = 2$, máme $0 \leq \tau_1 < \tau_2 \leq 1$ a interpolační polynom

$$p(r) = f(t_{n,1}, y(t_{n,1})) \frac{(t_{n+1} - r)}{h(\tau_2 - \tau_1)} + f(t_{n,2}, y(t_{n,2})) \frac{(r - t_n)}{h(\tau_2 - \tau_1)} \quad (4.14)$$

4.1. IMPLICITNÍ RUNGE-KUTTOVY METODY

Výpočtem (4.10) a (4.11) s interpolačním polynomem (4.14) získáváme Butcherovu tabulku

$$\begin{array}{c|cc} \tau_1 & \frac{\tau_2^2 - (\tau_2 - \tau_1)^2}{2(\tau_2 - \tau_1)} & -\frac{\tau_1^2}{2(\tau_2 - \tau_1)} \\ \tau_2 & \frac{\tau_2^2}{2(\tau_2 - \tau_1)} & \frac{(\tau_2 - \tau_1)^2 - \tau_1^2}{2(\tau_2 - \tau_1)} \\ \hline & \frac{\tau_2^2 - (1 - \tau_2)^2}{2(\tau_2 - \tau_1)} & \frac{(1 - \tau_1)^2 - \tau_1^2}{2(\tau_2 - \tau_1)} \end{array}$$

Pokud zvolíme $\tau_1 = 0$ a $\tau_2 = \frac{1}{2}$ tak dostaneme Butcherovu tabulku

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

což nám dává lichoběžníkovou metodu.

Pokud zvolíme $\tau_1 = \frac{1}{2} - \frac{1}{6}\sqrt{3}$, $\tau_2 = \frac{1}{2} + \frac{1}{6}\sqrt{3}$, což jsou koeficienty vycházející z Gaussovy kvadraturní formule viz [8], tak bude mít tabulka následující tvar.

$$\begin{array}{c|cc} \frac{3-\sqrt{3}}{6} & \frac{1}{2} & \frac{3-2\sqrt{3}}{12} \\ \frac{3+\sqrt{3}}{6} & \frac{3+2\sqrt{3}}{12} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$

Tato metoda se nazývá Gaussova dvoustupňová metoda.

Pokud zvolíme $\tau_1 = \frac{1}{3}$ a $\tau_2 = 1$, což jsou koeficienty vycházející z Radauovy pravé kvadraturní formule viz [8], získáme tabulku

$$\begin{array}{c|cc} \frac{1}{3} & \frac{5}{12} & \frac{-1}{12} \\ 1 & \frac{3}{4} & \frac{1}{4} \\ \hline & \frac{3}{4} & \frac{1}{4} \end{array}$$

a metodu, které říkáme Radau IIA dvoustupňová metoda.

Vyjmenované metody jsou jenom ukázkou těch, které můžeme nalézt. Vyjdeme-li z Radauovy levé kvadraturní formule získáme Radau IA metody, pro Lobattovu kvadraturní formuli získáváme Lobattovy metody apod.

Z metod vyšších stupňů uvedeme jen třístupňovou Radau IIA metodu, jejíž Butcherova tabulka je následující

$$\begin{array}{c|ccc} \frac{4-\sqrt{6}}{10} & \frac{88-7\sqrt{6}}{360} & \frac{296-196\sqrt{6}}{1800} & \frac{-2+3\sqrt{6}}{225} \\ \frac{4+\sqrt{6}}{10} & \frac{296+196\sqrt{6}}{1800} & \frac{88+7\sqrt{6}}{360} & \frac{-2-3\sqrt{6}}{225} \\ 1 & \frac{16-\sqrt{6}}{36} & \frac{16+\sqrt{6}}{36} & \frac{1}{9} \\ \hline & \frac{16-\sqrt{6}}{36} & \frac{16+\sqrt{6}}{36} & \frac{1}{9} \end{array}$$

Poznámka: Pravá strana rovnice (4.12) obsahuje všechny hodnoty $k_{n,j}$ pro $j = 1, \dots, s$ a tedy i hodnotu $k_{n,i}$. Pro vyjádření řešení tedy potřebujeme vyřešit soustavu nelineárních rovnic, která obsahuje s rovnic o s neznámých pro každý časový krok metody.

U každé numerické metody nás samozřejmě zajímá řád konvergence. Pro výše odvozené IRK metody byly J.C. Butcherem odvozeny tzv. zjednodušující vztahy [5]

$$\begin{aligned} B(p) : \quad & \sum_{i=1}^s b_i c_i^{k-1} = \frac{1}{k}, \quad k = 1, 2, \dots, p \\ C(q) : \quad & \sum_{j=1}^s a_{ij} c_j^{k-1} = \frac{c_i^k}{k}, \quad k = 1, 2, \dots, q, \quad i = 1, 2, \dots, s \\ D(r) : \quad & \sum_{i=1}^s b_i c_i^{k-1} a_{ij} = \frac{b_j}{k} (1 - c_j^k), \quad k = 1, 2, \dots, r, \quad j = 1, 2, \dots, s, \end{aligned}$$

jejichž vliv na řád metody je shrnut v následující větě.

Věta 4.1. [2] *Jestliže Runge-Kuttova metoda splňuje podmínky $B(p), C(q)$ a $D(r)$ pro $p \leq q + r + 1$ a $p \leq 2q + 2$ pak její řád konvergence je p .*

Jak je ukázáno v [2], tak Gaussova s -stupňová metoda je řádu $p = 2s$ a Radauova s -stupňová metoda je řádu $p = 2s - 1$. Tedy výše uvedná 2-stupňová Gaussova metoda je čtvrtého řádu, Radauova 2-stupňová metoda je třetího řádu a Radauova 3-stupňová metoda je řádu 5.

Dalším důležitým pojmem u každé metody je její stabilita. Nebudeme zde odvozovat stabilitu jednotlivých metod, toto lze nalézt např. v [2], ale uvedeme jen výsledky.

Definice 4.1. Pro lineární úlohu jsou RK metody *A-stabilní* pokud $|R(\eta)| < 1$ pro $\forall \eta \in \mathbb{C}$ s $Re(\eta) < 0$, kde $R(\eta) = 1 + \eta b^T (I - \eta A)^{-1} e$ a $e^T = (1, 1, \dots, 1)$ je s -rozměrný vektor.

Všechny Gaussovy, Radau IIA metody jsou pro lineární úlohy *A-stabilní*.

Definice 4.2. Numerickou metodu nazveme *L-stabilní*, jestliže je *A-stabilní* a navíc splňuje

$$\lim_{z \rightarrow \infty} R(z) = 0$$

Tvrzení 4.2. *Jestliže IRK metoda s regulární maticí A splňuje jednu z následujících podmínek*

$$a_{sj} = b_j, \quad j = 1, 2, \dots, s, \quad (4.15)$$

$$a_{i1} = b_1, \quad i = 1, 2, \dots, s \quad (4.16)$$

*potom $R(\infty) = 0$ a metoda je *L-stabilní*.*

Důkaz. Viz [8]. □

Metody, které splňují podmínku (4.15) se v anglické literatuře nazývají *stiffly accurate* a využívají se při řešení DAE. Těmi jsou Radau IIA nebo třeba Lobattovy metody.

Pro nelineární úlohy je vhodnější jiný typ stability a to *B-stabilita*. Její definici lze nalézt v [8]. V praxi je obtížné určit *B-stabilitu* z definice, proto se využívá následující tvrzení.

4.1. IMPLICITNÍ RUNGE-KUTTOVY METODY

Tvrzení 4.3. Runge-Kuttova metoda je B-stabilní jestliže

$$b_i \leq 0,$$

pro všechna $i = 1, \dots, s$ a

$$M = [b_i a_{ij} + b_j a_{ji} - b_i b_j]_{i,j=1}^s$$

je pozitivně definitní.

Důkaz. Viz [8]. □

Jestliže je metoda B-stabilní pak je i A-stabilní. Většina důležitých metod jako jsou Gaussovy, Radau IA a Radau IIA metody jsou B-stabilní.

Poznámka Při řešení tuhých problémů se setkáváme s problémem, že řád konvergence pro tuhý problém je nižší než pro netuhý problém. Tomuto jevu se říká redukce řádu. Více se o tomto lze dočíst např. v [1, 2].

IRK pro DAE s indexem 2

Implicitní Runge-Kuttovy metody se využívají pro řešení DAE a to zejména tzv. stiffly accurate metody, kterými jsou například Radau IIA a Lobattovy metody. První zmiňované jsou použity v kódu E. Hairera RADAU5 [8].

Při numerickém řešení systémů s indexem dva uvažujeme Hessenbergův tvar (2.19). Předpokládáme tedy, že $g_y f_z$ je regulární matice pro přesné řešení.

RK metody pro problém indexu 2 mají tvar

$$\begin{aligned} y_{n,i} &= y_n + h \sum_{j=1}^s a_{i,j} f(y_{n,j}, z_{n,j}) \quad i = 1, 2, \dots, s, \\ z_{n,i} &= z_n + h \sum_{j=1}^s a_{ij} l_{n,j}, \quad i = 1, 2, \dots, s, \\ y_{n+1} &= y_n + h \sum_{j=1}^s b_j f(y_{n,j}, z_{n,j}), \\ z_{n+1} &= z_n + h \sum_{j=1}^s b_j l_{n,j}, \\ 0 &= g(y_{n,i}), \quad i = 1, 2, \dots, s. \end{aligned}$$

Všimněme si, že máme navíc proměnnou $l_{n,i}$, která je potřeba pro vyřešení rovnice $g(y_{n,i}) = 0$. Získáme systém nelineárních rovnic, které můžeme vyřešit Newtonovou metodou nebo některou z jejích modifikací.

Poznámka: Při řešení DAE s indexem 2 Runge-Kuttovými metodami se často liší řád konvergence pro y_n a z_n . Například pro s-stupňovou Radau IIA metodu je řád pro y_n $p = 2s - 1$, zatímco pro z_n pouze $p = s$, více lze nalézt třeba v [8].

4.2. Metody zpětného derivování

Metody zpětného derivování (zkráceně BDF podle backward differentiation formulas) patří mezi implicitní lineární vícekrokové metody používané pro řešení tuhých problémů. Pro k -krokovou metodu zpětného derivování použijeme zkrácený zápis BDF $_k$. BDF opět nejprve odvodíme pro ODR a ukážeme jejich vlastnosti a poté přejdeme k jejich použití pro DAE indexu 2.

Mějme obyčejnou diferenciální rovnici

$$y' = f(t, y) \quad (4.17)$$

s počáteční podmínkou

$$y(t_0) = y_0.$$

k -krokovou BDF metodu získáme tak, že do (4.17) dosadíme t_{n+1} za t

$$y'(t_{n+1}) = f(t_{n+1}, y(t_{n+1}))$$

a nahradíme derivaci $y'(t_{n+1})$ derivací $P'_k(t_{n+1})$ interpolačního polynomu $P_k(t)$ stupně k , který prochází body $[t_{n+1}, y(t_{n+1})], [t_n, y(t_n)], \dots, [t_{n-k+1}, y(t_{n-k+1})]$. Tedy

$$P_k(t) = \sum_{j=-1}^{k-1} y(t_{n-j}) l_{j,n}(t),$$

kde $l_{j,n}(t)$, $j = -1, \dots, k-1$ jsou fundamentální polynomy Lagrangeova interpolačního polynomu, viz (4.9), pro uzly $t_{n+1}, \dots, t_{n-k+1}$. Použijeme-li

$$P'_k(t_{n+1}) \approx y'(t_{n+1}) = f(t_{n+1}, y(t_{n+1})),$$

tak spojením předchozích dvou vztahů dostáváme

$$\sum_{j=0}^k \alpha_j y(t_{n-j+1}) \approx h\beta f(t_{n+1}, y(t_{n+1})). \quad (4.18)$$

Nahradíme-li nyní v (4.18) přesné hodnoty $y(t_i)$, jejich aproximacemi y_i , $i = n-k+1, \dots, n+1$, tak k -kroková BDF metoda je dána předpisem

$$\sum_{j=0}^k \alpha_j y_{n+1-j} = h\beta f(t_{n+1}, y_{n+1}).$$

Koeficienty α_j, β lze pro konstantní délku kroku jednoduše vypočítat z následujících vztahů

$$\beta_k = \left(\sum_{n=1}^k \frac{1}{n} \right)^{-1}$$

$$P_k = \beta_k \sum_{n=1}^k \frac{1}{n} y^{k-n} (y-1)^n = \sum_{j=0}^k \alpha_j y^{k-j}.$$

4.2. METODY ZPĚTNÉHO DERIVOVÁNÍ

Např. pro $k = 3$

$$\beta_3 = \frac{1}{1 + \frac{1}{2} + \frac{1}{3}} = \frac{6}{11}$$

$$P_3 = \frac{6}{11}(y^2(y-1) + \frac{1}{2}y(y-1)^2 + \frac{1}{3}(y-1)^3) = y^3 - \frac{18}{11}y^2 + \frac{9}{11}y - \frac{2}{11}.$$

Pro případ $k = 1$ dostáváme implicitní Eulerovu metodu, která je zároveň i implicitní 1-stupňovou IRK metodou. Souhrnně jsou koeficienty α_j, β pro $k = 1, \dots, 6$ uvedeny v následující tabulce

k	β	α_0	α_1	α_2	α_3	α_4	α_5	α_6
1	1	1	-1					
2	$\frac{2}{3}$	1	$-\frac{4}{3}$	$\frac{1}{3}$				
3	$\frac{6}{11}$	1	$-\frac{18}{11}$	$\frac{9}{11}$	$-\frac{2}{11}$			
4	$\frac{12}{25}$	1	$-\frac{48}{25}$	$\frac{36}{25}$	$-\frac{16}{25}$	$\frac{3}{25}$		
5	$\frac{60}{137}$	1	$-\frac{300}{137}$	$\frac{300}{137}$	$-\frac{200}{137}$	$\frac{75}{137}$	$-\frac{12}{137}$	
6	$\frac{60}{147}$	1	$-\frac{360}{147}$	$\frac{450}{147}$	$-\frac{400}{147}$	$\frac{225}{147}$	$-\frac{72}{147}$	$\frac{10}{147}$

Samozřejmě i pro BDF metody nás bude zajímat jejich konvergence a stabilita. Pro $k \leq 6$ jsou BDF metody řádu k [6].

Pro lineární více krokové metody je typická D-stabilita.

Definice 4.3. Řekneme, že lineární více kroková metoda je *D-stabilní*, jestliže všechny kořeny prvního charakteristického polynomu

$$\rho(\xi) = \xi^k + \alpha_1 \xi^{k-1} + \dots + \alpha_{k-1} \xi + \alpha_k$$

leží uvnitř jednotkového kruhu $|z| \leq 1$ komplexní roviny \mathbb{C} a pokud některý kořen leží na hranici $|z| = 1$, pak je jednoduchý.

Všechny BDF k metody pro $k \leq 6$ jsou D-stabilní, viz [6].

Definice 4.4. Řekneme, že metoda je $A(\alpha)$ -stabilní, $\alpha \in (0, \frac{\pi}{2})$ jestliže oblast absolutní stability S obsahuje nekonečný klín

$$W_\alpha = \{re^{i\varphi} \in \mathbb{C} \mid r > 0, |\varphi - \pi| < \alpha\}.$$

Velikosti úhlů α lze nalézt např. v [8]. Protože pro BDF6 je úhel α už jen 18° , tak se tato metoda obvykle moc nepoužívá a pro $k \geq 7$ se tyto metody nepoužívají vůbec.

BDF pro DAE s indexem 2

BDF metody byly první metody použité pro výpočet řešení DAE a to C. W. Gearem v článku [7]. Velmi dobře pracují pro DAE s indexem 1, ale lze je použít i na semi-implicitní tvar DAE s indexem 2. BDF metody jsou základem např. kódu DASSL, viz [1].

Při řešení DAE s indexem 2 používáme BDF metodu na její Hessenbergův tvar indexu 2. Můžeme tedy předpokládat, že $g_y f_z$ je regulární, pak k -kroková BDF metoda má pro (2.19) předpis:

$$\begin{aligned} \sum_{j=0}^k \alpha_j y_{n+1-j} &= h \beta f(y_{n+1}, z_{n+1}), \\ 0 &= g(y_{n+1}). \end{aligned}$$

Jestliže je f vektor velikosti n a g vektor velikosti m , pak získáváme soustavu $n + m$ nelineárních rovnic pro $n + m$ proměnných. K jejímu vyřešení můžeme použít Newtonovu metodu nebo některou z jejích variant.

Pro lineární vícekrokové metody platí následující věta o existenci řešení.

Věta 4.4. *Předpokládejme, že pro řešení $y(t), z(t)$ (2.19) startovací hodnoty splňují*

$$y_j - y(t_j) = \mathcal{O}(h), \quad z_j - z(t_j) = \mathcal{O}(h), \quad g(y_j) = \mathcal{O}(h^2)$$

pro $j = 0, \dots, k-1$ a $t_j = t_0 + jh$. Jestliže $g_y f_z$ je regulární v okolí tohoto řešení a jestliže $\beta_k \neq 0$ potom nelineární systém

$$\begin{aligned} \sum_{i=0}^k \alpha_i y_i &= h \sum_{i=0}^k \beta_i f(y_i, z_i) \\ 0 &= g(y_k) \end{aligned}$$

má řešení pro $h \leq h_0$. Toto řešení je lokálně jednoznačné a splňuje

$$y_k - y(t_k) = \mathcal{O}(h), \quad z_k - z(t_k) = \mathcal{O}(h).$$

Důkaz. Viz [8]. □

Konvergence BDF metod pro DAE je shrnuta v následující větě

Věta 4.5. *Uvažujme problém indexu 2 (2.19), pro který platí, že $g_y f_z$ je regulární. Potom BDF k metoda má řád konvergence $p = k$, $k \leq 6$ tj.,*

$$y_j - y(t_j) = \mathcal{O}(h^{p+1}), \quad z_n - z(t_n) = \mathcal{O}(h^p), \quad \text{pro } t_n = nh \leq \text{konst},$$

vždy když počáteční podmínky splňují

$$y_j - y(t_j) = \mathcal{O}(h^{p+1}) \quad \text{pro } j = 0, \dots, k-1.$$

Důkaz. Viz [8]. □

Vidíme tedy, že i pro BDF je konvergence u algebraické proměnné nižší než u diferenciální proměnné.

5. Numerické řešení

V této kapitole budou numericky vyřešeny dvě soustavy DAE indexu 2 v Hessenbergově tvaru. První soustava je z příkladu 2.3, jedná se tedy o lineární soustavu DAE s nekonzstantními koeficienty a druhá soustava je soustava nelineárních DAE z příkladu 2.4. Obě soustavy budou řešeny pomocí kódu RADAU5, který napsali E. Hairer a G. Wanner. Originální kód je napsán v jazyce FORTRAN, ale my jsme použili jeho přepis od D. Bichsela do MATLABu. Tyto kódy lze stáhnout z <http://www.unige.ch/~hairer/software.html>.

Kód RADAU5 je určený pro řešení počátečních úloh prvního řádu tuhých diferenciálních rovnic a pro DAE ve tvaru

$$My' = f(y, t),$$

kde M je diagonální konstantní matice, kterou nazýváme matice hmotnosti (z anglického mass matrix). Pro ODR je diagonála matice M nenulová. Pro DAE jsou na hlavní diagonále 1 pro diferenciální proměnné a 0 pro algebraické proměnné, tedy DAE je v semi-explicitním tvaru.

Program RADAU5 je založen na implicitních Runge-Kuttových metodách, konkrétně využívá Radau IIA metody stupňů 3, 5 a 7. Řád metody, kterou má být daný problém počítán, si tedy lze zvolit. Navíc je v programu implementována korekce časového kroku, pro dosažení předepsané tolerance. V RADAU5 jsou pro tuto časovou korekci naprogramovány dva postupy, standardní kontrola kroku a prediktivní kontrola kroku, více se lze dočíst v [2].

My jsme v našem řešení využili pouze 3 stupňové Radau IIA metody, která je uvedena v kapitole 4.1. Jak již bylo řečeno, tato metoda má řád konvergence 5.

Vstupy kódu jsou soustava rovnic (soubory `priklad2_3fce.m` a `priklad2_4fce.m` na přiloženém CD), interval, na kterém chceme soustavu řešit, a pro DAE matice hmotnosti (soubory `priklad2_3Mass.m` a `priklad2_4Mass.m`) a index soustavy. Lze zadat i Jakobián soustavy (soubory `priklad2_3Jac.m` a `priklad2_4Jac.m`). Jestliže zadán není, je počítán numericky, což zvýší čas potřebný k výpočtu. V našem případě se jednalo o velmi malý časový rozdíl, téměř nepostřehnutelný. Výpočty lze jednoduše spustit ze skriptů `priklad2_3` a `priklad2_4`.

Příklad 1

Máme soustavu lineární DAE s nekonzstantními koeficienty z příkladu 2.3:

$$\begin{aligned}y_1' &= \left(\alpha - \frac{1}{2-t}\right)y_1 + (2-t)\alpha z + \frac{3-t}{2-t}e^t \\y_2' &= \frac{1-\alpha}{t-2}y_1 - y_2 + (\alpha-1)z + 2e^t \\0 &= (t+2)y_1 + (t^2-4)y_2 - (t^2+t-2)e^t\end{aligned}\tag{5.1}$$

kde α je parametr s počátečními podmínkami

$$y_1(0) = y_2(0) = 1.$$

Známe tedy pouze počáteční hodnoty diferenciálních proměnných y , pro výpočet ovšem musíme zadat i počáteční hodnotu algebraické proměnné z . Tu můžeme vypočítat ze vztahu

$$g_y f(y(0), z(0), 0) = 0. \quad (5.2)$$

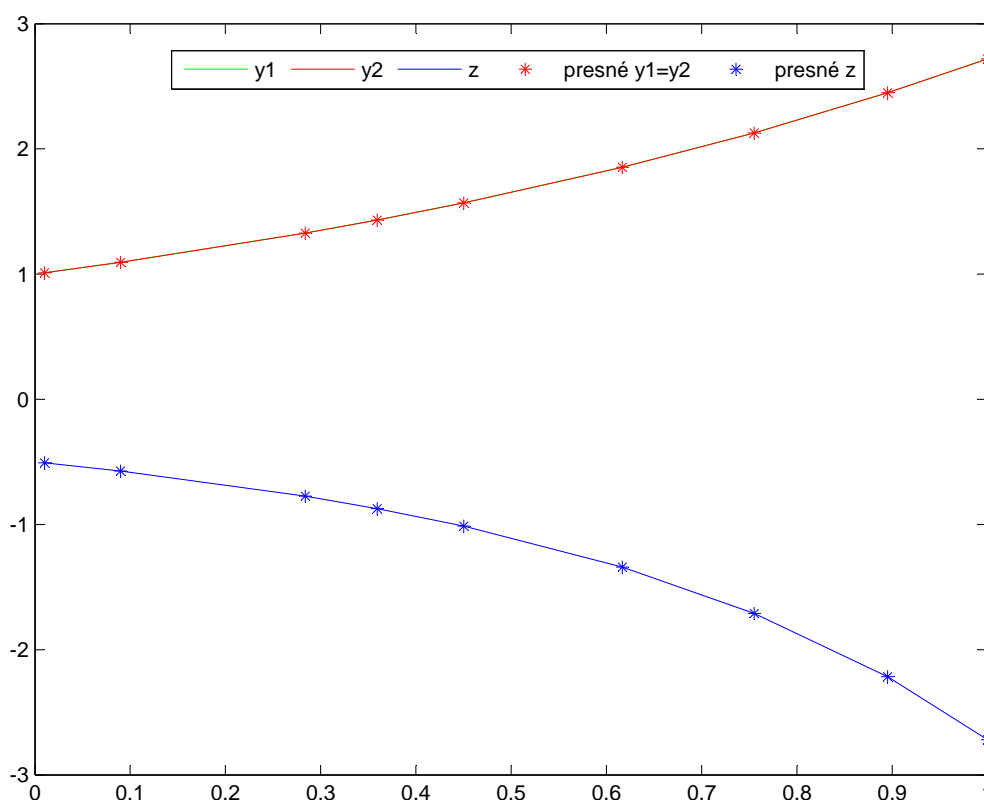
Jednoduchým výpočtem získáme $z(0) = -\frac{1}{2}$.

Vypočítané numerické řešení samozřejmě srovnáme s přesným řešením

$$y_1 = y_2 = e^t, \quad z = -\frac{e^t}{2-t}.$$

Soustavu jsme řešili na intervalu $\langle 0, 1 \rangle$ a se zadaným Jakobiánem.

Pro různé hodnoty parametru α se řešení výrazně neliší, řádově asi 10^{-6} . Rozdíl v různých hodnotách α se projevil v počtu potřebných časových kroků. Pro malé hodnoty α bylo řešení spočítáno pomocí několika málo kroků (např. pro $\alpha = 2$ bylo potřeba 11 kroků), kdežto pro $\alpha = 100$ už bylo potřeba kroků téměř 500. Řešení soustavy pro $\alpha = 1$ je znázorněno na obrázku 5.1, kde plné čáry značí numericky vypočtené řešení pomocí kódu RADAU5 a symbol $*$ značí přesné řešení v uzlech numerického řešení, a je z nich patrná proměnlivá délka kroku.



Obrázek 5.1: Řešení soustavy (2.15) pro $\alpha = 1$.

Příklad 2

Máme soustavu nelineární DAE z příkladu 2.4

$$\begin{aligned}y_1' &= y_1 y_2^2 z^2 \\y_2' &= y_1^2 y_2^2 - 3y_2^2 z \\0 &= y_1^2 y_2 - 1.\end{aligned}\tag{5.3}$$

s počátečními podmínkami

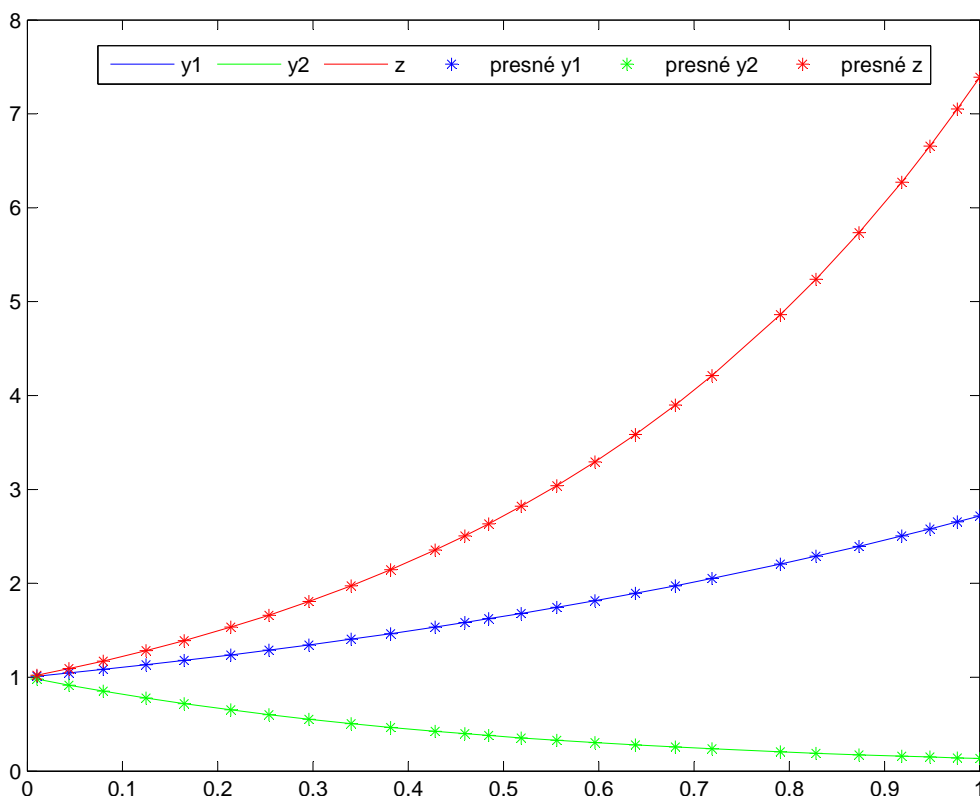
$$y_1(0) = 1, y_2(0) = 1.$$

Počáteční hodnotu z opět vyjádříme ze vztahu (5.2) a dostaneme $z = 1$.

I tuto soustavu jsme řešili na intervalu $\langle 0, 1 \rangle$ a se zadaným Jakobiánem. Řešení je znázorněno na obrázku 5.2. Výsledek opět srovnáme s přesným řešením

$$y_1(t) = e^t, y_2(t) = e^{-2t}, z(t) = e^{2t},$$

které je, stejně jako v předchozím příkladu, vypočítáno v uzlech numerického řešení a zobrazeno pomocí symbolu *. Plné čáry značí řešení vypočtené numericky pomocí kódu RADAU5.



Obrázek 5.2: Řešení soustavy (2.15).

Z obrázků je vidět, že pro výpočet nelineární soustavy 5.3 bylo na stejně dlouhém intervalu použito většího počtu časových kroků než v případě lineární soustavy s konstantními koeficienty 5.1, což je očekávaný výsledek.

6. Závěr

Cílem práce bylo rozdělení algebraicko-diferenciálních rovnic podle indexu a popsat jejich numerické řešení pro index 2. Nejprve byly DAE popsány teoreticky, byly ukázány jejich základní vlastnosti a odlišnosti od obyčejných diferenciálních rovnic. Dále pak bylo uvedeno několik typů indexů s důrazem na nejpoužívanější, jimiž jsou diferenciální index a Kroneckerův index. Při numerickém řešení byla uvažována počáteční úloha.

Bylo ukázáno, že DAE mají blízko k tuhým ODR, proto jsme se zaměřili na přímé diskretizační metody původně určené pro řešení tuhých problémů. Odvozeny byly implicitní Runge-Kuttovy metody a metody zpětného derivování. Poté bylo ukázáno jak tyto metody použít pro Hessenbergovy tvary DAE s indexem 2 a jaké problémy s tím souvisí.

U obou skupin metod je přesnost pro algebraickou proměnnou menší než pro diferenciální. Obecně se u IRK vyskytuje redukce řádu, proto se pro tuhé systémy a tedy i pro DAE používají stiffly accurate metody, u kterých tento jev není tolik citelný.

Závěrem byl proveden výpočet dvou soustav DAE pomocí 3-stupňové Radau IIA metody implementované v kódu RADAU5. Výsledky těchto dvou výpočtů byly srovnány i s přesným řešením soustav. Vypočtená řešení můžeme považovat za velmi dobré aproximace řešení přesného.

Práce nabízí spoustu možností, jak ji doplnit a rozšířit například o řešení DAE s indexem 1 nebo naopak DAE vyšších řádů. Dále by bylo možné popsat specializované metody pro DAE s indexem 2 a metody redukce indexu. V rámci teoretického popisu by bylo možné přidat např. kanonické tvary.

Literatura

- [1] ASCHER, Uri M. a Linda R. PETZOLD. *Computer methods for ordinary differential equations and differential-algebraic equations*. 1. Philadelphia: Society for Industrial and Applied Mathematics, 1998. ISBN 08-987-1412-5.
- [2] ATKINSON, Kendall E. *Numerical solution of ordinary differential equations*. 1. Hoboken: Wiley, 2009. ISBN 978-0-470-04294-6.
- [3] BAUMANN, Sascha. *Coupled electromagnetic field circuit simulation: modeling and numerical analysis*. Berlin: Logos, 2012. ISBN 978-383-2531-911.
- [4] BRENAN, Kathryn E., S. L. CAMPBELL a Linda R. PETZOLD. *Numerical solution of initial-value problems in differential-algebraic equations*. New York: North-Holland, 1989. ISBN 04-440-1511-6.
- [5] BUTCHER, J. C. Implicit Runge-Kutta processes. *Mathematics of Computation*. 1964, **18**(85), 50-64. DOI: 10.1090/S0025-5718-1964-0159424-9. ISSN 0025-5718. Dostupné také z: <http://www.ams.org/jourcgi/jour-getitem?pii=S0025-5718-1964-0159424-9>
- [6] ČERMÁK, Libor. *Numerické metody pro řešení diferenciálních rovnic*. Brno: Litera Brno, 2013. ISBN 978-80-903586-7-6.
- [7] GEAR, C. Simultaneous Numerical Solution of Differential-Algebraic Equations. *IEEE Transactions on Circuit Theory*. 1971, **18**(1), 89-95. DOI: 10.1109/TCT.1971.1083221. ISSN 0018-9324. Dostupné také z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1083221>
- [8] HAIRER, E. a Gerhard WANNER. *Solving ordinary differential equations II*. 2nd rev. ed. Heidelberg: Springer, 2010. ISBN 978-364-2052-217.
- [9] KUNKEL, Peter a Volker MEHRMANN. *Differential-algebraic equations: analysis and numerical solution*. Zürich: European Mathematical Society Publishing House, 2006. ISBN 978-303-7190-173.
- [10] LAMOUR, René, Roswitha MÄRZ a Caren TISCHENDORF. *Differential-algebraic equations: a projector based analysis*. New York: Springer, 2013. Differential-algebraic equations forum. ISBN 36-422-7554-0.
- [11] Index Concepts for Differential-Algebraic Equations. MEHRMANN, Volker. *Encyclopedia of applied and computational mathematics*. 1. Berlin, Heidelberg: Springer, 2015, s. 676-681. ISBN 978-3-540-70529-1.
- [12] RIAZA, Ricardo. *Differential-algebraic systems analytical aspects and circuit applications*. 1. Singapore, SG: World Scientific, 2008. ISBN 978-981-2791-818.
- [13] SIMERSKÁ, Carmen. Systémy algebro-diferenciálních rovnic. *Pokroky matematiky, fyziky a astronomie*. 2005, **50**(3), 182-192. Dostupné také z: <http://dml.cz/dmlcz/141270>