

MC SYLLABUS 15.3

P.A. BEENTJES

K. DEKKER

P.W. HEMKER

M. VAN VELDHIJZEN

**COLLOQUIUM STIJVE
DIFFERENTIAALVERGELIJKINGEN**

DEEL 3

MATHEMATISCH CENTRUM

AMSTERDAM 1975

AMS(MOS) subject classification scheme (1970): 65L05, 92-00, 92-04

ISBN 90 6196 118 1

INHOUD

VOORWOORD

- | | |
|---|-----|
| 10. Schets van een convergentietheorie voor discretizaties van stijve differentiaalvergelijkingen | 221 |
| M. van Veldhuizen, Math. Inst. Rijksuniversiteit, Utrecht | |
| 11. Testresultaten van methoden voor stijve vergelijkingen | 247 |
| P.A. Beentjes, K. Dekker, Math. Centrum Amsterdam | |
| 12. Testresultaten van methoden met een stapkeuze-strategie | 298 |
| P.W. Hemker, Math. Centrum, Amsterdam | |

VOORWOORD

Het colloquium "stijve differentiaalvergelijkingen" werd van september 1971 tot december 1972 aan het Mathematisch Centrum gehouden en had tot doel een overzicht te geven van recent onderzoek dat verricht werd op het gebied van het numeriek oplossen van "stijve" differentiaalvergelijkingen.

De eerste twee delen van het colloquiumverslag bevatten onder andere een historisch overzicht, eenstaps- en meerstapsmethoden en toepassingen uit de biomathematica.

In het eerste hoofdstuk van deze laatste syllabus wordt een schets van een convergentietheorie gegeven. De resterende hoofdstukken zijn gewijd aan een bespreking van testresultaten van de verschillende methoden die in de vorige twee deeltjes van dit colloquiumverslag zijn behandeld.

P.A. Beentjes

10. Schets van een convergentietheorie voor discretisaties van stijve differentiaalvergelijkingen

10.1. Inleiding

Een overzicht zal worden gegeven van enkele resultaten die verkregen zijn in de loop van een onderzoek dat er op gericht was inzicht te verkrijgen in de wijze waarop een oplossing verkregen met behulp van een discretisatiemethode, de oplossing van een stijve differentiaalvergelijking benadert. Het blijkt dat de situatie essentieel ingewikkelder is dan voor een gewone niet stijve differentiaalvergelijking. Belangrijk is dat onder redelijk algemene omstandigheden de oplossing van een stijve differentiaalvergelijking gesplitst kan worden in een "langzaam variërende" (reguliere) component en een "snel dalende" (singuliere) component. Evenzo voor de ge-discretiseerde vergelijkingen. Een foutenanalyse zal dan de volgende drie vragen moeten beantwoorden:

- I Hoe benadert de reguliere component van de differentiaalvergelijking een "daarbij passende" reguliere component van de discretisatie?
- II Hoe benadert de singuliere component van de differentiaalvergelijking een "daarbij passende" singuliere component van de discretisatie?
- III Gegeven een startvector, welke reguliere en singuliere component hoort hierbij voor de differentiaalvergelijking en welke reguliere en singuliere component hoort hierbij voor de discretisatie?

We verduidelijken een en ander aan een zeer eenvoudig voorbeeld. Bezie voor $t \geq 0$ de differentiaalvergelijking in de $\mathbb{R}^2$:

$$(10.1.1) \quad \frac{dx}{dt} = \begin{bmatrix} -1000 & 0 \\ e^{-t} & 0 \end{bmatrix} x .$$

Iedere reguliere component is een veelvoud van

$$(10.1.2) \quad x_R(t) \equiv \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

en iedere singuliere component is een veelvoud van

$$(10.1.3) \quad x_S(t) \equiv \bar{D} \begin{bmatrix} e^{-1000t} \\ e^{-1001t} \\ 1001 \end{bmatrix} .$$

Immers, zowel x_R als x_S zijn oplossingen van de differentiaalvergelijking. Bovendien zijn x_R en x_S als functies lineair onafhankelijk. Het is echter niet duidelijk waarom $x_R + 10^{-12}x_S$ niet als reguliere component wordt gekozen. Anders gezegd, is het mogelijk het begrip reguliere component te definiëren, zonder het heuristische "langzame variëren"? Dit is mogelijk en dat zal straks kort geschetst worden. Analoog de singuliere component.

Als discretisatie kiezen we voor ons voorbeeld de A-stabiele trapeziumregel. Als x_i een benadering voor $x(t_i)$, $t_i = ih$, $i = 0, 1, 2, \dots$ voorstelt, h de stapgrootte, dan na enig schrijfwerk:

$$(10.1.4) \quad x_{i+1} = \begin{bmatrix} \frac{1-500h}{1+500h} & 0 \\ \frac{1-500h}{1+500h} e^{-t_i-h} + e^{-t_i} & \frac{h}{2} \\ & 1 \end{bmatrix} x_i .$$

Iedere reguliere component is veelvoud van

$$(10.1.5) \quad (x_R)_i \equiv \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad i = 0, 1, 2, \dots$$

en iedere singuliere component is veelvoud van

$$(10.1.6) \quad (x_S)_i \equiv \begin{bmatrix} p^i \\ \frac{h}{2} \frac{pe^{-h}+1}{pe^{-h}-1} e^{-t_i} p^i \end{bmatrix}, \quad i = 0, 1, 2, \dots$$

$$\text{met } p = \frac{1-500h}{1+500h} .$$

We zullen zeggen dat $\{(x_R)_i\}_{i \geq 0}$, (daarmee is de rij van vectoren $(x_R)_i$ bedoeld, $i = 0, 1, 2, \dots$) past bij $x_R(t)$ als $x_R(0) = (x_R)_0$.

En $\{(x_S)_i\}_{i \geq 0}$ past bij $x_S(t)$ als

eerste coörd. $(x_S)_0 = \text{eerste coörd. } x_S(0)$.

De reden waarom we dit zo doen wordt later duidelijk (in 10.4).

We kunnen nu de vragen I, II, en III beantwoorden.

I: Kennelijk $x_R(t_i) = (x_R)_i$, dus bij elkaar passende reguliere componenten zijn identiek in de steunpunten $t_0, t_1, \dots$.

II: Dit is lastiger. p^i benadert zeker niet goed $e^{-1000} t_i$, tenminste als $500h \geq 1$. Toch zal men vaak de trapeziumregel wensen te gebruiken met een stapgrootte zodanig dat (in ons geval) $500h \gg 1$. Dan is de discrete singuliere component dus een slechte benadering voor de continue singuliere component.

III: Dit vereist wat rekenwerk.

Zij $x_0 = \begin{bmatrix} \delta \\ 1 \end{bmatrix}$, δ een in modulus klein reëel getal. $\delta \neq 0$ kan optreden ten gevolge van afrondfouten, of omdat we nog een klein restje van het sneldalende stuk over hebben.

Met x_0 associëren we de oplossing $x(t)$ van de differentiaalvergelijking met $x(0) = x_0$. Stel

$$x(t) = \alpha x_R(t) + \beta x_S(t).$$

Dan, speciaal voor $t = 0$:

$$\begin{array}{ccc} \beta = \delta & & \beta = \delta \\ & \Rightarrow & \\ \alpha - \frac{\beta}{1001} = 1 & & \alpha = 1 - \frac{\delta}{1001}. \end{array}$$

Met x_0 associëren we ook de oplossing $\{x\}_{i \geq 0}$ van de discretisatie bij stapgrootte h . Stel:

$$x_i = \alpha_h (x_R)_i + \beta_h (x_S)_i.$$

Speciaal $i = 0$ beziend, vinden we

$$\alpha_h = 1 - \delta \frac{h}{2} \frac{pe^{-h} + 1}{pe^{-h} - 1}, \quad \beta_h = \delta.$$

Merk op $\beta_h = \beta$ en $\alpha \neq \alpha_h$. We vinden

$$x(t_i) - x_i = \delta\{x_S(t_i) - (x_S)_i\} + (\alpha - \alpha_h)(x_R)_i.$$

Voor $h = 1$

$$\max_{i \in \mathbb{N}} \|x(t_i) - x_i\|_{\infty} \approx 1.16\delta.$$

Maar voor $h = 10$

$$\max_{i \in \mathbb{N}} \|x(t_i) - x_i\|_{\infty} \approx 3.5\delta.$$

Deze fout is dus, samenvattend, het gevolg van het feit dat de startvector op verschillende wijze ontbonden wordt in startvectoren voor de reguliere en singuliere component voor de beide situaties: het continue geval (namelijk de differentiaalvergelijking) en haar discretisatie.

In het nu volgende zullen we het voorgaande voor meer algemene situaties generaliseren. Eerst echter zullen we een manier moeten vinden om reguliere en singuliere componenten op een niet heuristische wijze te definiëren. We zullen ons beperken tot een lineaire inhomogene differentiaalvergelijking in de $\mathbb{R}^2$. Hieraan zijn de belangrijkste verschijnselen die we theoretisch kunnen voorspellen goed aantoonbaar. We zullen ons op de beantwoording van vraag I richten. In een kort tijdsbestek lijkt het ondoenlijk de andere vragen (II en III) begrijpelijk te analyseren en te beantwoorden.

10.2. Definitie en existentie van reguliere componenten

Bezie de differentiaalvergelijking:

$$(10.2.1) \quad \frac{dx}{dt} = A(t)x$$

$x \in \mathbb{R}^2$, $A(t)$ de 2×2 matrix

$$(10.2.2) \quad A(t) = \begin{bmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{bmatrix}.$$

Aanname (II₁) $A(t)$ is gedefinieerd voor alle $t \in \mathbb{R}$. $A(t)$ is voldoende vaak continu differentieerbaar en de matrix coëfficiënten voldoen voor alle

$t \in \mathbb{R}$ aan:

$$a_{11}(t) < -m_{11} < m_{22} \leq a_{22}(t) \leq M_{22}$$

$$|a_{12}(t)| \leq m_{12}, \quad a_{21}(t) \leq m_{21}.$$

Voorts:

$$m_{11} > 0, \quad M_{22} < m_{11}.$$

In het nu volgende is het nuttig om steeds te bedenken dat $|a_{12}|$ en $|a_{21}|$ "klein" zijn ten opzichte van $|a_{11}|$. Merk op, dat als $|a_{12}(t)| \equiv |a_{21}(t)| \equiv 0$ voor alle $t \in \mathbb{R}$, dat dan op triviale wijze de eerste coördinaat van de oplossing de sneldalende component voorstelt (immers $a_{11}(t) < 0$) en dat dan de tweede coördinaat de langzaam variërende component voorstelt.

Neem nu aan, heuristisch sprekend, dat ook als $a_{12} \neq 0$, $a_{21} \neq 0$ toch zoiets als een langzaam variërende component bestaat en ook een sneldalende component. Zij voor $t = t_0$ de startvector $x(t_0) \in \mathbb{R}^2$ gegeven. Deze startvector bepaalt een oplossing $x(t)$ van (10.2.1). We mogen verwachten dat $x(t)$ voor $t \gg t_0$ nog uitsluitend de reguliere component bevat. Willen we voor $t \geq 0$ de reguliere component vinden dan zullen we dus t_0 sterk negatief moeten kiezen. Met behulp van de variatie-van-constanten formule hebben we:

$$x_1(t) = \phi_1(t)\phi_1^{-1}(t_0)x_1(t_0) + \int_{t_0}^t \phi_1(t)\phi_1^{-1}(s)a_{12}(s)x_2(s)ds$$

$$x_2(t) = \phi_2(t)\phi_2^{-1}(t_0)x_2(t_0) + \int_{t_0}^t \phi_2(t)\phi_2^{-1}(s)a_{11}(s)x_2(s)ds,$$

waarin ϕ_1 en ϕ_2 gedefinieerd zijn door:

$$\phi_1(t) \stackrel{\text{D}}{=} \exp\left[\int_0^t a_{11}(r)dr\right],$$

$$\phi_2(t) \stackrel{\text{D}}{=} \exp\left[\int_0^t a_{22}(r)dr\right].$$

Dus $\phi_1(t)\phi_1^{-1}(s) = \exp\left[\int_s^t a_{11}(r)dr\right]$ en voor a_{11} constant is dit $\exp[(t-s)a_{11}]$.

Als $t_0 \rightarrow -\infty$ dan $\phi_1(t)\phi_1^{-1}(t_0)x_1(t_0) \rightarrow 0$ als x een langzaam variërende component is, althans dit verwachten we. Op heuristische gronden is het ook niet onmogelijk dat

$$\left| \int_{t_0}^t \phi_1(t)\phi_1^{-1}(s)a_{12}(s)x_2(s)ds \right| < \infty \quad \text{voor } t_0 \rightarrow -\infty.$$

$$\begin{aligned} \text{Omdat } x_2(t) &= \phi_2(t)\phi_2^{-1}(t_0)x_2(t_0) + \int_{t_0}^t \phi_2(t)\phi_2^{-1}(s)a_{21}(s)x_2(s)ds \\ &= \phi_2(t)[\phi_2^{-1}(t_0)x(t_0) + \int_{t_0}^0 \phi_2^{-1}(s)a_{21}(s)x_1(s)ds] \\ &\quad + \int_0^t \phi_2(t)\phi_2^{-1}(s)a_{21}(s)x_1(s)ds \\ &= \phi_2(t)x_2(0) + \int_0^t \phi_2(t)\phi_2^{-1}(s)a_{21}(s)x_1(s)ds \end{aligned}$$

zullen we de vergelijkingen

$$(10.2.3) \quad x_1(t) = \int_{-\infty}^t \phi_1(t)\phi_1^{-1}(s)a_{12}(s)x_2(s)ds$$

$$x_2(t) = \phi_2(t)x_2(0) + \int_0^t \phi_2(t)\phi_2^{-1}(s)a_{21}(s)x_1(s)ds$$

onderzoeken. Met behulp van (10.2.3) zullen we trachten reguliere oplossingen te definiëren op een niet heuristische wijze en hun bestaan aan te tonen. In feite werken we iets algemener. Als we de differentiaalvergelijking:

$$(10.2.4) \quad \frac{dx}{dt} = A(t)x + b(t)$$

hebben, b gedefinieerd voor alle $t \in \mathbb{R}$, zullen we bezien:

$$x_1(t) = \int_{-\infty}^t \phi_1(t) \phi_1^{-1}(s) [a_{12}(s)x_2(s) + b_1(s)] ds$$

(10.2.5)

$$x_2(t) = \phi_2(t)x_2(0) + \int_0^t \phi_2(t) \phi_2^{-1}(s) [a_{21}(s)x_1(s) + b_2(s)] ds.$$

We vervolgen nu op een niet-heuristische wijze.

Aanname (II2) De eigenwaarden van de matrix

$$(10.2.6) \quad \begin{bmatrix} -m_{11} & m_{12} \\ -m_{21} & m_{22} \end{bmatrix}$$

zijn reëel en verschillend. De grootste noemen we μ_2 , de kleinste μ_1 .

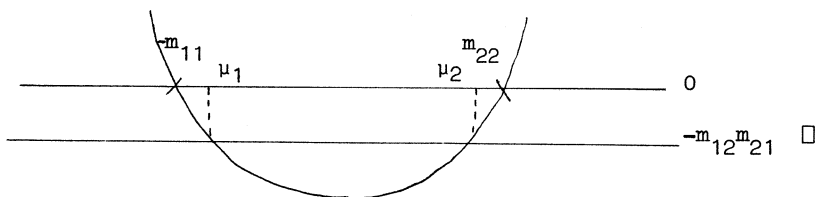
Lemma (II3) $-m_{11} \leq \mu_1 < \mu_2 \leq m_{22}$

Bewijs Het karakteristieke polynoom is

$$\lambda^2 + (m_{11} - m_{22})\lambda - m_{11}m_{22} + m_{12}m_{21} = 0.$$

De discriminant D is positief. Zie nu de figuur:

karakteristiek polynoom



In de $\mathbb{R}^2$ geeft $\|x\| = |x_1| + |x_2|$ de 1-norm van een vector $x \in \mathbb{R}^2$ aan.

Afspraak (II4) Voor een $\tilde{v} \in (\mu_1, \mu_2)$ is $\tilde{C}_{\tilde{v}}$ de ruimte van de continue afbeeldingen x van $(-\infty, 0]$ in de $\mathbb{R}^2$, die nog voldoen aan

$$\sup_{t \leq 0} \frac{\|x(t)\|}{e^{\tilde{v}t}} < \infty.$$

Op $\tilde{C}_{\tilde{v}}$ is $\|\cdot\|_{\tilde{v}}$ een norm, voor $x \in \tilde{C}_{\tilde{v}}$ gedefinieerd door

$$\|x\|_{\tilde{v}} = \sup_{t \leq 0} \frac{\|x(t)\|}{e^{\tilde{v}t}}.$$

Kennelijk is $\tilde{C}_{\tilde{v}}$ een Banachruimte.

Stelling (II5) Neem aan dat b , beperkt tot $(-\infty, 0]$ in $\tilde{C}_{\tilde{v}}$. Dan bezit (10.2.5) voor iedere $x_2(0)$ een unieke oplossing x in $\tilde{C}_{\tilde{v}}$. Deze oplossing is ook de enige oplossing van (10.2.4) in $\tilde{C}_{\tilde{v}}$, waarvan in $t = 0$ de tweede coördinaat $x_2(0)$ is. Deze oplossing kan worden afgeschat door de oplossing r van

$$\begin{aligned} (10.2.7) \quad r_1(t) &= \int_{-\infty}^t e^{-m_{11}(t-s)} [m_{12}r_2(s) + |b_1(s)|] ds \\ r_2(t) &= e^{m_{22}t} |x_2(0)| + \int_t^0 e^{m_{22}(t-s)} [m_{21}r_1(s) + |b_2(s)|] ds \end{aligned}$$

voor $t \in (-\infty, 0]$. Dus $|x_i(t)| \leq r_i(t)$ voor $i = 1, 2$, $t \leq 0$.

(Het voorgaande impliceert dat alle oneigenlijke integralen voor $x \in \tilde{C}_{\tilde{v}}$ bestaan).

We bewijzen St.(II5) niet; dat zou te veel werk zijn. We proberen te begrijpen wat de conclusies van St. (II5) betekenen. De eerste twee conclusies van St.(II5) spreken voor zich. We knopen aan bij de vergelijking (10.2.7). (10.2.7) noemen we een majorant voor (10.2.5). Unieke oplosbaarheid van (10.2.7) is verzekerd door de eerste twee conclusies van St.(II5).

Lemma (II6) De oplossing r van (10.2.7) voldoet aan

$$(10.2.8) \quad \frac{dr}{dt} = \begin{bmatrix} -m_{11} & m_{12} \\ -m_{21} & m_{22} \end{bmatrix} r + \begin{bmatrix} |b_1(t)| \\ -|b_2(t)| \end{bmatrix}.$$

Bewijs: Door differentiëren. $\square$

Oplossingen van (10.2.8) kunnen we aangeven. Daartoe hebben we de eigenvectoren van (10.2.6) nodig. Zij e_1 de eigenvector behorend bij μ_1 , e_2 de eigenvector behorend bij μ_2 .

$$\text{Dan geldt } e_1 = (1, \xi_2)^T = (1, \frac{m_{21}}{m_{22}-\mu_1})^T$$

$$\text{en } e_2 = (\xi_1, 1)^T = (\frac{m_{12}}{m_{11}+\mu_2}, 1)^T.$$

Zij S de matrix

$$(10.2.9) \quad S = \begin{bmatrix} 1 & \frac{m_{12}}{m_{11}+\mu_2} \\ \frac{m_{21}}{m_{22}-\mu_1} & 1 \end{bmatrix}.$$

Lemma (II7) $\det S \neq 0$

Bewijs Wegens $(m_{22}-\mu_1)(-m_{11}-\mu_1) + m_{12}m_{21} = 0$ (karakteristiek polynoom van (10.2.6)) geldt:

$$\frac{m_{12}}{m_{11}+\mu_2} \cdot \frac{m_{21}}{m_{22}-\mu_1} = \frac{m_{11}+\mu_1}{m_{11}+\mu_2}$$

$$\mu_1 + m_{11} = \frac{m_{11} + m_{22} - \sqrt{D}}{2}, \quad D \text{ de discriminant van het karakteristiek polynoom en}$$

$$m_{11} + \mu_2 = \frac{m_{11} + m_{22} + \sqrt{D}}{2}. \quad \text{Uit } D > 0 \text{ volgt nu}$$

$$\det S = 1 - \frac{m_{12}m_{21}}{(m_{11}+\mu_2)(m_{22}-\mu_1)} > 0. \quad \square$$

Zij

$$(10.2.10) \quad d(t) = S^{-1} \begin{bmatrix} |b_1(t)| \\ -|b_2(t)| \end{bmatrix},$$

dan is (10.2.8) equivalent met:

$$(10.2.11) \quad \frac{dy}{dt} = \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix} y + d(t) .$$

Een oplossing van (10.2.11) is

$$(10.2.12) \quad \begin{aligned} y_1(t) &= \int_{-\infty}^t e^{\mu_1(t-s)} d_1(s) ds \\ y_2(t) &= e^{\mu_2 t} y_2(0) + \int_0^t e^{\mu_2(t-s)} d_2(s) ds. \end{aligned}$$

Dit is de enige oplossing van (10.2.11) in $\tilde{C}_v$ want iedere andere oplossing met dezelfde $y_2(0)$ scheelt $\alpha (e^{\mu_1 t}; 0)^T$ met $y(t)$. En $(e^{\mu_1 t}, 0)^T \notin \tilde{C}_v$.

Wegens de eerste twee conclusies van St.(II5) geldt dan:

Lemma (II8) Als $y_2(0)$ zo gekozen wordt dat $(Sy(0))_2 = |x_2(0)|$, dan is $Sy(t)$ (y volgens (10.2.12)) de unieke oplossing van (10.2.8) in $\tilde{C}_v$.

Gevolg De beide coördinaten van $Sy(t)$ zijn voor alle $t \leq 0$ niet-negatief.

Immers volgens St.(II5) geldt voor de oplossing x van (10.2.5):

$$|x_i(t)| \leq r_i(t) \quad i = 1, 2, \quad t \leq 0 .$$

Voor $t \geq 0$ geldt een analogon van St.(II5). Ter voorbereiding bezien we de eigenwaarden van de matrix

$$(10.2.13) \quad \begin{bmatrix} -m_{11} & m_{12} \\ m_{21} & M_{22} \end{bmatrix} .$$

Het karakteristiek polynoom van (2.13) is:

$$\lambda^2 + (m_{11} - M_{22})\lambda - m_{11}M_{22} - m_{12}m_{21} = 0$$

met discriminant

$$D = (m_{11} - M_{22})^2 + 4m_{11}M_{22} + 4m_{12}m_{21} .$$

Dus

$$D = (m_{11} + M_{22})^2 + 4m_{12}m_{21} > 0$$

onder onze aannamen. (Dat was aanname (II1)).

Er zijn dus twee reële eigenwaarden.

Afspraak (II9) Zij $v \in \mathbb{R}$, v groter dan de grootste eigenwaarde van (10.2.13). Dan is C_v de ruimte der op $[0, \infty)$ gedefinieerde continue functies x waarvoor

$$\sup_{t \geq 0} \frac{\|x(t)\|}{e^{vt}} < \infty.$$

Op C_v is $\|\cdot\|_v$ een norm, voor $x \in C_v$, gedefinieerd door

$$\|x\|_v \equiv \sup_{t \geq 0} \frac{\|x(t)\|}{e^{vt}}.$$

Kennelijk is C_v een Banachruimte.

Stelling (II10) De differentiaalvergelijking (10.2.4) bezit voor iedere startvector $x_0 \in \mathbb{R}^2$ op het tijdstip 0 een unieke oplossing in C_v , mits $b \in C_v$. Bovendien wordt de oplossing puntsgewijs en coördinaatsgewijs gemajoreerd door de oplossing van de differentiaalvergelijking:

$$(10.2.14) \quad \frac{dr}{dt} = \begin{bmatrix} -m_{11} & m_{12} \\ m_{21} & M_{22} \end{bmatrix} r + \begin{bmatrix} |b_1(t)| \\ |b_2(t)| \end{bmatrix}$$

$$\text{met} \quad r(0) = \begin{bmatrix} |(x_0)_1| \\ |(x_0)_2| \end{bmatrix}.$$

We bewijzen St.(II10) niet. De eerste conclusie is overigens een zeer wel bekend resultaat; alleen de tweede conclusie schijnt nieuw te zijn en is dan een uitbreiding van het lemma van Gronwall. (10.2.14) zullen we weer een majorant noemen. Oplossingen van de majorant kunnen bepaald worden,

omdat de oplossing van (10.2.14) expliciet is aan te geven in termen van de eigenwaarden en eigenvectoren van (10.2.13) en de inhomogene term. We laten dit hier achterwege.

St.(II10) wordt toegepast in combinatie met St.(II5). Voor gegeven $x_2(0)$ bepaalt St.(II5) een oplossing x van (10.2.4) op $(-\infty, 0]$. Daarmee is ook $x(0)$ bekend en met deze $x(0)$, dus verkregen m.b.v. St.(II5), wordt St.(II10) toegepast. Op deze wijze verkrijgen we een oplossing van (10.2.4) op de gehele $\mathbb{R}$, die bepaald wordt door de gegeven $x_2(0)$, een 2^e coördinaat in $t = 0$.

Oplossingen van (10.2.4) op deze manier verkregen noemen we *reguliere oplossingen*. Hiermee is dus de niet-heuristische definitie gegeven. Rest nog om in te zien dat de zo gedefiniëerde reguliere oplossing ook "langzaam varieert". Daartoe zullen we heel expliciet moeten zeggen wat dat wel is, dat "langzaam variëren". We hebben gekozen voor de volgende oplossing.

In plaats van (10.2.4) bezien we de differentiaalvergelijking:

$$(10.2.15) \quad \frac{dx}{dt} = \begin{bmatrix} \frac{a_{11}(t)}{\epsilon} & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{bmatrix} x + b(t) .$$

Dus a_{11} vervangen door $\frac{a_{11}}{\epsilon}$, $\epsilon > 0$. Neem aan dat St.(II5) en St.(II10) toepasbaar zijn op (10.2.15) voor alle $\epsilon \in (0, \epsilon_0]$. Hoe kleiner ϵ , hoe meer de differentiaalvergelijking (10.2.15) stijf is. "Langzaam variëren" wordt nu geïnterpreteerd in termen van de afgeleiden van de reguliere oplossing van (10.2.15).

Stelling (II11) Neem aan dat a_{11} , a_{12} , a_{21} en a_{22} N -keer continu differentieerbaar zijn voor alle $t \in \mathbb{R}$ en dat hun afgeleiden begrensd zijn, uniform op $\mathbb{R}$. Laat eveneens de eerste N afgeleiden van b bestaan en laat gelden dat deze afgeleiden, beperkt tot $(-\infty, 0]$, in $\tilde{C}_{\tilde{v}(\epsilon)}$ liggen en dat ze, beperkt tot $[0, \infty)$, in $C_{v(\epsilon)}$ liggen. Dan geldt dat iedere reguliere oplossing $x_\epsilon(t)$ van (10.2.15) N -keer continu differentieerbaar is en voor $k = 1, 2, \dots, N$:

$$(i) \quad \frac{d^k x_\epsilon}{dt^k}, \text{ beperkt tot } (-\infty, 0], \text{ ligt in } \tilde{C}_{\tilde{v}(\epsilon)} \quad \text{en}$$

$$\left\| \frac{d^k x_\epsilon}{dt^k} \right\|_{\tilde{v}(\epsilon)} \text{ is begrensd uniform in } \epsilon \in (0, \epsilon_0].$$

(ii) $\frac{d^k x_\epsilon}{dt^k}$ beperkt tot $[0, \infty)$, ligt in $C_{v(\epsilon)}$ en

$$\left\| \frac{d^k x_\epsilon}{dt^k} \right\|_{v(\epsilon)} \text{ is begrensd uniform in } \epsilon \in (0, \epsilon_0].$$

Stelling (II11) wordt niet bewezen, hoewel ze een direkt gevolg van St.(II5) en St.(II10) is. St.(II11) beschrijft inderdaad ons heuristische begrip "langzaam variëren". Wat we namelijk van plan zijn, vooruitlopend op een volgende paragraaf, is het onderzoek van convergentie van discretisatie methoden. Zoals bekend neemt daarbij de discretisatiefout een belangrijke plaats in. Deze discretisatiefout wordt bepaald door zekere afgeleiden van de exacte (analytische) oplossing. St.(II11) zegt dan dat de discretisatiefout voor een reguliere oplossing in de norm van $\tilde{C}_{\tilde{v}(\epsilon)}$ en $C_{v(\epsilon)}$ begrensd is onafhankelijk van ϵ , d.w.z. onafhankelijk van de sterk negatieve eigenwaarden van de stijve differentiaalvergelijking. Uit het voorbeeld gegeven in de inleiding, blijkt dat dit voor een niet reguliere oplossing niet geldt. In het genoemde voorbeeld was een oplossing (daar alvast heuristisch de singuliere component genoemd) (zie (10.1.3))

$$\begin{bmatrix} e^{-1000t} \\ e^{-1001t} \\ -\frac{e^{-1001t}}{1001} \end{bmatrix}.$$

Hier staat -1000 voor $\frac{a_{11}}{\epsilon}$; kies maar $a_{11} = -1$ en $\epsilon = \frac{1}{1000}$.

Afgeleiden zijn in norm van de orde van grootte

$$1000^k e^{-1000t}$$

voor de k^{de} afgeleide of meer algemeen

$$\frac{1}{\epsilon^k} e^{-1/\epsilon t}.$$

Met andere woorden, in het algemeen zal een willekeurige oplossing van (10.2.15) geen uniform in ϵ begrensde afgeleiden hebben. Dat geeft natuurlijk moeilijkheden bij de discretisatiefout voor kleine ϵ en daarom zijn reguliere oplossingen volgens St(II5) en St(II10) ingevoerd.

10.3. Reguliere oplossingen van differentievergelijkingen

Om in de volgende paragraaf de convergentietheorie te kunnen behandelen hebben we ook een definitie van een reguliere component van een recursie of differentievergelijking nodig. Daaraan is deze paragraaf gewijd. We behandelen uitsluitend één-staps-recursies in de $\mathbb{R}^2$.

Bezie de differentievergelijking:

$$(10.3.1) \quad x_{i+1} = A_i x_i + b_i$$

met

$$(10.3.2) \quad A_i = \begin{bmatrix} a_{11}^{(i)} & a_{12}^{(i)} \\ a_{21}^{(i)} & a_{22}^{(i)} \end{bmatrix} \quad \text{en} \quad b_i = \begin{bmatrix} (b_i)_1 \\ (b_i)_2 \end{bmatrix}.$$

Aanname (III1) A_i en b_i zijn gedefinieerd voor alle $i \in \mathbb{Z}$.
Bovendien geldt voor alle $i \in \mathbb{Z}$:

$$|a_{12}^{(i)}| \leq c_{12}, \quad |a_{21}^{(i)}| \leq c_{21},$$

$$|a_{11}^{(i)}| \leq c_{11}, \quad c_{11} < c_{22} \leq a_{22}^{(i)} \leq c_{22}.$$

Aanname (III2) De matrix

$$(10.3.3) \quad \begin{bmatrix} c_{11} & c_{12} \\ -c_{21} & c_{22} \end{bmatrix}$$

bezit twee verschillende reële eigenwaarden $\tau_1, \tau_2, \tau_1 < \tau_2$.

Volgens lemma (II3) geldt dan

$$c_{11} \leq \tau_1 < \tau_2 \leq c_{22}.$$

We kunnen nu vervolgen als in de vorige paragraaf. We vatten slechts de resultaten samen.

Afspraak (III3) $\tilde{D}_\rho$ is de ruimte van de oneindige rijtjes vectoren uit de $\mathbb{R}^2$:

$$\{x\}_{i \leq 0} = \{x_0, x_{-1}, x_{-2}, \dots\}$$

die voldoen aan

$$\sup_{i \in \mathbb{N}} \frac{\|x_{-i}\|}{\rho^{-i}} < \infty$$

($\mathbb{N}$ is de verzameling der niet negatieve gehele getallen).

Op $\tilde{D}_\rho$ is $\|\cdot\|_\rho$ een norm,

$$\|\{x\}_{i \leq 0}\|_\rho = \sup_{i \in \mathbb{N}} \frac{\|x_{-i}\|}{\rho^{-i}}$$

Afspraak (III4) D_ρ is de ruimte van de oneindige rijtjes vectoren uit de $\mathbb{R}^2$

$$\{x\}_{i \geq 0} = \{x_0, x_1, x_2, \dots\}$$

die voldoen aan

$$\sup_{i \in \mathbb{N}} \frac{\|x_i\|}{\rho^i} < \infty.$$

Op D_ρ is $\|\cdot\|_\rho$ een norm

$$\|\{x\}_{i \geq 0}\|_\rho = \sup_{i \in \mathbb{N}} \frac{\|x_i\|}{\rho^i}.$$

Stelling (III5) Laat ρ voldoen aan $\tau_1 < \rho < \tau_2$.

Dan is er precies één oplossing $\{x\}_{i \leq 0}$ van de recursie (10.3.1) in $\tilde{D}_\rho$ mits $\{b\}_{i \leq 0} \in \tilde{D}_\rho$ bij gegeven $(x_0)_2$ (dit is de 2^{de} coördinaat van x_0).

Puntsgewijs en coördinaatsgewijs wordt deze oplossing gemajoreerd door de oplossing $\{r\}_{i \leq 0}$ van de recursie

$$(10.3.4) \quad r_{i+1} = \begin{bmatrix} c_{11} & c_{12} \\ -c_{21} & c_{22} \end{bmatrix} r_i + \begin{bmatrix} |(b_i)_1| \\ |(b_i)_2| \end{bmatrix}$$

met $(r_0)_2 = |(x_0)_2|$.

De eisen leggen $\{r\}_{i \leq 0}$ éénduidig vast en $(r_{-i})_j \geq 0$ voor alle $i \in \mathbb{N}$ en $j = 1, 2$. Ook $(r_{-i})_j \geq |(x_{-i})_j|$ voor alle $i \in \mathbb{N}$ en $j = 1, 2$ (j duidt de coördinaat aan).

Zij ρ groter dan de grootste eigenwaarde van de matrix

$$(10.3.5) \quad \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}.$$

Stelling (III6) Bij iedere $x_0 \in \mathbb{R}^2$ is er precies één oplossing $\{x\}_{i \geq 0}$ van de recursie (10.3.1) in D_ρ mits $\{b\}_{i > 0} \in D_\rho$.

Puntsgewijs en coördinaatsgewijs wordt $\{x\}_{i \geq 0}$ gemajoreerd door de oplossing $\{r\}_{i \geq 0}$ van

$$(10.3.6) \quad r_{i+1} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} r_i + \begin{bmatrix} |(b_i)_1| \\ |(b_i)_2| \end{bmatrix}$$

met $r_0 = (|(x_0)_1|, |(x_0)_2|)^T$.

Bovendien $\{r\}_{i \geq 0} \in D_\rho$.

We noemen (10.3.4) en (10.3.6) majorerende recursies. Hun oplossingen kunnen op elementaire wijze berekend worden.

Bij gegeven $x_2(0)$ behoort volgens St.(III5) een oplossing van de recursie (10.3.1) voor $i \leq 0$; met behulp van St.(III6) breiden we deze oplossing uit tot een oplossing op $\mathbb{Z}$. Deze oplossing, afhankelijk van $x_2(0)$ noemen we *de reguliere component bij de startwaarde $x_2(0)$ voor $i = 0$* .

10.4. Convergentietheorie voor reguliere componenten.

We gaan uit van de differentiaalvergelijking in de $\mathbb{R}^2$:

$$(10.4.1) \quad \frac{dy}{dt} = A(t)y + g(t),$$

met A en g gedefinieerd voor alle $t \in \mathbb{R}$. A en g zullen in het algemeen nog van een parameter ε afhangen, $\varepsilon > 0$. Deze afhankelijkheid verwerken we niet in de notatie omdat dat de notatie maar zou vertroebelen. De afhankelijkheid van ε wordt wel altijd stilzwijgend verondersteld, tenzij uitdrukkelijk anders aangegeven.

Aanname (IV1) Voor alle $t \in \mathbb{R}$ is er een transformatie $L(t)$ (eventueel weer afhankelijk van ε) zo dat voor alle $t \in \mathbb{R}$ en $\varepsilon \in (0, \varepsilon_0]$:

$$(i) \quad L(t)A(t)L^{-1}(t) = A(t) = \begin{bmatrix} \frac{\lambda_1(t)}{\varepsilon} & 0 \\ 0 & \lambda_2(t) \end{bmatrix}.$$

Hierin zijn λ_1 en λ_2 eventueel nog afhankelijk van ε , maar $\frac{\lambda_1(t)}{\varepsilon} \rightarrow \infty$ voor $\varepsilon \rightarrow 0$.

$$(ii) \quad \frac{d}{dt} L(t) \cdot L^{-1}(t) \equiv B(t) \text{ bestaat.}$$

Met de substitutie

$$(10.4.2) \quad x(t) = L(t)y(t)$$

gaat (10.4.1) over in

$$(10.4.3) \quad \frac{dx}{dt} = [A(t) + B(t)]x + b(t)$$

met

$$(10.4.4) \quad b(t) = L(t)g(t).$$

Aanname (IV2) De differentiaalvergelijking (10.4.3) voldoet voor alle $\varepsilon \in (0, \varepsilon_0]$ aan de aannamen van stelling (II5) en (II10) (deze aannamen zijn o.a. aanname (II2) en enkele aannamen over het groeikarakter van b). Bovendien zij aan de voorwaarden van St.(II11) voldaan.

Bij differentiaalvergelijking (10.4.3) behoren dus de twee Banach-ruimten $\tilde{C}_{\tilde{v}(\varepsilon)}$ en $C_{v(\varepsilon)}$ (zie afspraak (II4) en (II9)). $\tilde{v}$ en v hangen in ons geval van ε af. Overigens kan men narekenen dat voor $\varepsilon \in (0, \varepsilon_0]$

$$\tilde{C}_{\tilde{v}(\varepsilon)} \subset \tilde{C}_{\tilde{v}(\varepsilon_0)} \text{ en } C_{v(\varepsilon)} \subset C_{v(\varepsilon_0)}.$$

Naast de differentiaalvergelijking hebben we een differentievergelijking. Zij h een gegeven stapgrootte. Laat $y_i \in \mathbb{R}^2$ behoren bij het tijdstip $t_i \stackrel{\text{D}}{=} ih$, $i \in \mathbb{Z}$. Een éénstapsmethode voor stijve differentiaalvergelijkingen geeft dan een recursie die aan y_i een y_{i+1} toevoegt voor $i \in \mathbb{Z}$. Stel nu

$$(10.4.5) \quad x_i = L(t_i)y_i.$$

Zo verkrijgen we een recursie in de $\{x_i\}$ die we noteren als in (10.3):

$$(10.4.6) \quad x_{i+1} = A_i x_i + b_i$$

(vergelijk (10.3.1), waarbij we moeten bedenken dat nu A_i en b_i afhangen van twee parameters, namelijk h en ϵ).

Voorbeeld (IV3) We discretiseren de differentiaalvergelijking (10.4.1) met behulp van een bekende methode (Liniger en Willoughby):

$$(10.4.7) \quad y_{i+1} = y_i + \alpha h A(t_i) y_i + \beta h A(t_{i+1}) y_{i+1} + h(\alpha g(t_i) + \beta g(t_{i+1})).$$

Gebruikmakend van aanname (IV1) en de substitutie (10.4.5) vinden we na enig rekenwerk (10.4.6) met

$$(10.4.8) \quad A_i = [1 - \beta h \Lambda(t_{i+1})]^{-1} [1 + \alpha h \Lambda(t_i)] + \\ + h [1 - \beta h \Lambda(t_{i+1})]^{-1} W_i(h) [1 + \alpha h \Lambda(t_i)]$$

en

$$(10.4.9) \quad b_i = h \{ \alpha [1 - \beta h \Lambda(t_{i+1})]^{-1} [1 + h W_i(h)] b(t_i) + \\ + \beta [1 - \beta h \Lambda(t_{i+1})]^{-1} b(t_{i+1}) \}$$

met $h W_i(h) = [L(t_{i+1}) - L(t_i)] L(t_i)^{-1}$.

In de convergentietheorie (volgens Dahlquist) speelt het begrip discretisatiefout een grote rol. Daarin treden afgeleiden van de exacte oplossing op en, zoals al in § 2 is getoond, kan dit problemen geven als we ons niet beperken tot reguliere componenten. We beperken ons dus tot reguliere componenten. We willen dan laten zien dat "bij elkaar passende" re-

guliere componenten van de differentiaalvergelijking en de differentievergelijking elkaar benaderen met een fout die in slechts geringe mate afhankelijk is van ϵ en die voor $\epsilon \in (0, \epsilon_0]$ te majoreren is door een van ϵ onafhankelijke majorant. Meer wiskundig gezegd, we zullen eisen dat uniform in ϵ , $\epsilon \in (0, \epsilon_0]$, het verschil tussen de componenten in een of andere norm naar nul gaat voor $h \rightarrow 0$. Het lijkt voor de hand te liggen om dan ook te eisen dat uniform in ϵ , $\epsilon \in (0, \epsilon_0]$ de discretisatiefout $\rightarrow 0$ voor $h \rightarrow 0$ in een of andere norm. Hier liggen wat problemen; het zou echter te ver voeren om die hier nader te precisieren. In ons 2×2 geval is de oplossing eenvoudig.

Definitie (IV4) Als x een reguliere oplossing is van de differentiaalvergelijking (10.4.3) dan is

$$\delta_i = \frac{1}{h} \{x(t_{i+1}) - A_i x(t_i) - b_i\}$$

de *R-discretisatiefout* voor de methode (10.4.6) in $t = t_i$. (R-discretisatiefout vanwege de beperking tot een reguliere component. De terminologie werd door dr. P.J. v.d. Houwen aan de auteur gesuggereerd).

Let op het verschil tussen de R-discretisatiefout en de discretisatiefout !

Zoals al eerder opgemerkt gold voor x en zijn afgeleiden dat ze beperkt tot $(-\infty, 0]$ element zijn van $\tilde{C}_v(\epsilon)$ en beperkt tot $[0, \infty)$ element van $C_v(\epsilon)$. Dus is de rij $\{x(t_i)\}_{i \in \mathbb{Z}}$ beperkt tot $i \leq 0$ element van

$$\tilde{D}_{\exp(h\tilde{v}(\epsilon))}$$

en beperkt tot $i \geq 0$ element van

$$D_{\exp(hv(\epsilon))}$$

(Vergelijk afspraak III3 en III4).

Dus geldt

$$\{\delta\}_{i \leq 0} \in \tilde{D}_{\exp(h\tilde{v}(\epsilon))}$$

en

$$\{\delta\}_{i \geq 0} \in D_{\exp(hv(\epsilon))} .$$

Definitie (IV5) De methode (10.4.6) heet *R-consistent* indien uniform in $\epsilon \in (0, \epsilon_0]$ geldt

$$\lim_{h \rightarrow 0} \max\{\|\{\delta\}_{i \leq 0}\| \exp(h\nu(\epsilon)), \|\{\delta\}_{i \geq 0}\| \exp(h\nu(\epsilon))\} = 0.$$

De methode heet *R-consistent en van de orde p* indien uniform in $\epsilon \in (0, \epsilon_0]$

$$\max\{\|\{\delta\}_{i \leq 0}\| \exp(h\nu(\epsilon)), \|\{\delta\}_{i \geq 0}\| \exp(h\nu(\epsilon))\} = O(h^p), \quad h \rightarrow 0.$$

Voorbeeld (IV6) Voor de methode (10.4.7) valt de orde, gedefinieerd volgens def(IV5), samen met de orde zoals gedefinieerd voor een niet-stijve differentiaalvergelijking. Immers, in (10.4.7) wordt de discretisatiefout bij substitutie van $y(t_i)$ in plaats van y_i , y een reguliere oplossing van (10.4.1) (y gedefinieerd als $L^{-1}x$, x een reguliere oplossing van (10.4.3)), coördinaatsgewijs in t_i gegeven door (cf. Liniger en Willoughby):

$$-h \int_0^1 (\theta - 1 - \alpha) y_j''(t_i + \theta h) d\theta$$

($j = 1, 2$) als $\alpha + \beta = 1$, $\alpha \neq \frac{1}{2}$ en

$$\frac{h^2}{4} \int_0^1 2\theta(\theta - 1) y_j'''(t_i + \theta h) d\theta$$

als $\alpha = \beta = \frac{1}{2}$.

Na transformatie en vermenigvuldiging met

$$[1 - \beta h \Lambda(t_{i+1})]^{-1}$$

waarvan de sup-norm $\|\cdot\|$ begrensd is uniform in $\epsilon \in (0, \epsilon_0]$ en h , vinden we δ_i en puntsgewijs (d.w.z. voor $h \rightarrow 0$ en $ih \rightarrow t$, t het punt dat we bezien):

$$\begin{aligned} \|\delta_i\| &= O(h) & \alpha \neq \frac{1}{2} \quad \alpha + \beta &= 1 \\ \|\delta_i\| &= O(h^2) & \alpha &= \beta = \frac{1}{2} \end{aligned}.$$

Voorts nog (wegens St.(II11))

$$\{\delta\}_{i \leq 0} \in \tilde{D}_{\exp(h\nu(\epsilon))}, \quad \{\delta\}_{i \geq 0} \in D_{\exp(h\nu(\epsilon))}$$

waaruit het beweerde volgt.

Het is niet waar dat het begrip orde gedefinieerd in def(IV5) samenvalt met het gewone begrip orde. Zie het volgende voorbeeld.

Voorbeeld (IV7)

Beschouw de differentiaalvergelijking

$$\frac{dx}{dt} = A(t)x.$$

Hier is A onafhankelijk van ϵ (dus niet "stiff").

Methode:

$$x_{i+1} = [1 - \frac{1}{2}hA(t_{i+1})]^{-1} x_i + \frac{1}{2}hA(t_i)[1 - \frac{1}{2}hA(t_{i+1})]^{-1} x_i.$$

Of ook:

$$x_{i+1} = x_i + \frac{1}{2}hA(t_i)x_i + \frac{1}{2}hA(t_{i+1})x_{i+1} + \frac{h^2}{4} [A(t_{i+1})A(t_i) - A(t_i)A(t_{i+1})] [1 - \frac{1}{2}hA(t_{i+1})]^{-1} x_i.$$

Dus is de methode een trapeziumregel met een stoorterm die $O(h^3)$ is want

$$A(t_{i+1})A(t_i) - A(t_i)A(t_{i+1}) = O(h).$$

Dus is de methode van orde 2 in de gebruikelijke zin. Niet echter als de differentiaalvergelijking stijf is. Zij A als in (10.2.15), dus (A nu weer afhankelijk van ϵ):

$$A(t) = \begin{bmatrix} \frac{a_{11}(t)}{\epsilon} & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{bmatrix}$$

waarbij de coëfficiënten aan zulke voorwaarden voldoen dat St.(II5), St.(II10) en St.(II11) toepasbaar zijn. Voor een reguliere oplossing x is de R-discretisatiefout in t_i :

$$\delta_i = \frac{1}{h} \{x(t_{i+1}) - [1 - \frac{1}{2}hA(t_{i+1})]^{-1} [1 + \frac{1}{2}hA(t_i)]x(t_i) +$$

$$+ \frac{h^2}{4} [1 - \frac{1}{2}hA(t_{i+1})]^{-1} [A(t_{i+1})A(t_i) - A(t_i)A(t_{i+1})] [1 - \frac{1}{2}hA(t_{i+1})]^{-1} x(t_i) \} .$$

Dus is $\delta_i(h)$ de R-discretisatiefout van de trapeziumregel plus een stoorterm. Nu geldt:

$$\begin{aligned} & A(t_{i+1})A(t_i) - A(t_i)A(t_{i+1}) = \\ = & \begin{bmatrix} [a_{12}, a_{21}]_{t_i}^{t_{i+1}} & \frac{1}{\epsilon} [a_{11}, a_{12}]_{t_i}^{t_{i+1}} + [a_{21}, a_{22}]_{t_i}^{t_{i+1}} \\ \frac{1}{\epsilon} [a_{21}, a_{11}]_{t_i}^{t_{i+1}} + [a_{22}, a_{21}]_{t_i}^{t_{i+1}} & -[a_{12}, a_{21}]_{t_i}^{t_{i+1}} \end{bmatrix} \end{aligned}$$

waarin voor functies f en g gedefinieerd is:

$$[f, g]_{t_i}^{t_{i+1}} \stackrel{D}{=} f(t_{i+1})g(t_i) - f(t_i)g(t_{i+1}) .$$

Dus in het algemeen $|A(t_{i+1})A(t_i) - A(t_i)A(t_{i+1})| \neq O(h)$ uniform in ϵ voor $h \rightarrow 0$. Men kan narekenen (door de inverse van $I - \frac{1}{2}hA(t_{i+1})$ op te schrijven) dat de gehele stoorterm $O(h^2)$ is, uniform in ϵ . Dus is de orde volgens def(IV5) niet 2 maar hoogstens 1.

Toegegeven, het voorbeeld is wat artificieel. Men kan echter laten zien dat als de 3^e orde semi-implicite methode (van der Houwen) R-consistent is, dat dan de orde hooguit 2 is in de zin van IV5. Soms zijn er ook problemen met de polynoommethoden (van der Houwen) in Runga-Kutta-vorm. Het kost echter veel werk om dat te laten zien. Daarom laten we dit hier achterwege.

Na de consistentie behandelen we de stabiliteit. Zoals in (10.3) kort vermeldt is, behoorden bij een recursie (10.4.6) onder zekere aannamen twee ruimten $\tilde{D}_{\tilde{\rho}}$ en D_{ρ} . Hier hangen, voor een differentievergelijking gebracht op de vorm (10.4.6), ρ en $\tilde{\rho}$ af van h en ϵ , dus $\rho(h, \epsilon)$ en $\tilde{\rho}(h, \epsilon)$.

Aanname (IV8) Voor alle $\epsilon \in (0, \epsilon_0]$ en voor alle $h \in (0, h_0(\epsilon)]$ zijn de ruimten $\tilde{D}_{\tilde{\rho}(h, \epsilon)}$ en $D_{\rho(h, \epsilon)}$ gedefinieerd bij de recursie (10.4.6) volgens de afspraken (III3) en (III4) en St.(III5) en St.(III6). Bovendien nemen we aan dat $\tilde{\rho}(h, \epsilon)$ en $\rho(h, \epsilon)$ zo gekozen kunnen worden voor alle $\epsilon \in (0, \epsilon_0]$ en $h \in (0, h_0(\epsilon)]$ dat

$$\tilde{D}_{\exp(h\tilde{\nu}(\epsilon))} \subset \tilde{D}_{\tilde{\rho}(h,\epsilon)}$$

en

$$D_{\exp(h\nu(\epsilon))} \subset D_{\rho(h,\epsilon)}$$

(Voldoende hiervoor is, $\exp(h\tilde{\nu}(\epsilon)) \geq \tilde{\rho}(h,\epsilon)$ en $\exp(h\nu(\epsilon)) \leq \rho(h,\epsilon)$).

Zij x een reguliere oplossing van de getransformeerde differentiaalvergelijking (10.4.3). Zij $e_i = x(t_i) - x_i$, $\{x\}_{i \in \mathbb{Z}}$ een reguliere oplossing van de recursie (getransformeerd) (10.4.6). Dan geldt dus

$$\{e\}_{i \leq 0} \in \tilde{D}_{\tilde{\rho}(h,\epsilon)} \text{ en } \{e\}_{i \geq 0} \in D_{\rho(h,\epsilon)}$$

Wegens:

$$x(t_{i+1}) - A_i x(t_i) - b_i = h\delta_i$$

en

$$x_{i+1} - A_i x_i - b_i = 0$$

voldoet $\{e\}_{i \in \mathbb{Z}}$ aan:

$$(10.4.10) \quad e_{i+1} - A_i e_i = h\delta_i$$

Dus is $\{e_i\}_{i \in \mathbb{Z}}$ wegens St.(III5) en (III6) een reguliere oplossing van de recursie (10.4.10).

Definitie (IV9) De methode (10.4.6) heet *R-stabiel* indien voor alle $\epsilon \in (0, \epsilon_0]$ en alle $h \in (0, h_0(\epsilon)]$ geldt:

$$\|\{e\}_{i \in \mathbb{Z}}\|_{h,\epsilon} \leq C_1 |(e_0)_2| + C_2 \|\{\delta\}_{i \in \mathbb{Z}}\|_{h,\epsilon}$$

waarin

$$\|\{e\}_{i \in \mathbb{Z}}\|_{h,\epsilon} = \max(\max_{i \leq 0} \|e_i\|_{\tilde{\rho}(h,\epsilon)}, \max_{i \geq 0} \|e_i\|_{\rho(h,\epsilon)})$$

en C_1 en C_2 constanten onafhankelijk van h en ϵ zijn.

Definitie (IV10) De methode (10.4.6) heet *R-convergent* indien uniform in $\epsilon \in (0, \epsilon_0]$ geldt:

$$\lim_{h \rightarrow 0} (e_2) = 0$$

impliceert

$$\lim_{h \rightarrow 0} \|\{e\}_{i \in \mathbb{Z}}\|_{h, \epsilon} = 0.$$

Stelling (IV11)

R-consistentie $\wedge$ R-stabiliteit $\Rightarrow$ R-convergentie.

R-consistentie van de orde $p \wedge$ R-stabiliteit impliceren

$$\|\{e\}_{i \in \mathbb{Z}}\|_{h, \epsilon} = O(h^p) \quad \text{uniform in } \epsilon$$

indien $(e_2)_0 = O(h^p)$ uniform in ϵ .

Bewijs Door toepassing van de definities. $\square$

Wegens St.(IV11) doet een R-stabiele en R-consistente discretisatie precies wat we verwachten. Merk op dat de uitspraak van St.(IV11) een wiskundige vertaling is van "convergentie afhankelijk van de stijfheid", iets wat kennelijk alleen voor reguliere componenten geldt. St.(IV11) beantwoordt dus vraag I van de inleiding.

Opmerking (IV12) De functie $h_0(\epsilon)$ hangt sterk af van de stabiliteitseigenschappen van de methode. Zo zal in het algemeen voor de expliciete methoden van Van der Houwen $h_0(\epsilon) \rightarrow 0$ als $\epsilon \rightarrow 0$. $h_0(\epsilon) = h_0 = \text{constant}$ zal dus betekenen dat R-stabiliteit optreedt voor alle $\epsilon \in (0, \epsilon_0]$ en $h \in (0, h_0]$. Dit mogen we verwachten als we een impliciete methode gebruiken met sterke "dampings eigenschappen". Zo bijvoorbeeld de gefitte methode (10.4.7). Maar bijvoorbeeld niet de trapeziumregel (hoewel beide A-stabiel zijn).

Opmerking (IV13) Het voorgaande is geldig voor differentiaalvergelijkingen die gedefinieerd zijn voor alle $t \in \mathbb{R}$. Dit lijkt een grote beperking, te meer daar nogal sterke eisen voor alle $t \in \mathbb{R}$ worden opgelegd. Men kan echter bewijzen dat het voldoende is de eisen op te leggen voor $t \in [a, b]$ met $[0, T] \subset (a, b)$ als men geïnteresseerd is in de oplossing op $[0, T]$.

Men verandert namelijk buiten $[a, b]$ de differentiaalvergelijking zo, dat aan alle eisen is voldaan. En men kan bewijzen dat dit op de uiteindelijke foutenanalyse, waarin de vragen I, II en III worden beantwoord, geen essentiële invloed heeft, omdat de totale fout op $[0, T]$ onafhankelijk is van de uitbreiding of verandering buiten $[a, b]$.

Opmerking (IV14) Het vereist nogal wat rekenwerk om R-stabiliteit aan te tonen voor een gegeven methode. We doen dat daarom niet maar we merken wel op (zonder bewijs) dat bijvoorbeeld de methode (10.4.7) met $0 \leq \alpha \leq \frac{1}{4}$ altijd R-stabiel is voor constante $h_0(\epsilon) = h_0$ met h_0 en ϵ_0 voldoende klein. Hetzelfde geldt voor de (expliciete) Taylor-methode

$$y_{i+1} = y_i + h\dot{y}_i + \alpha h^2 \ddot{y}_i$$

en ook hier is h_0 onafhankelijk van ϵ indien de methode toegepast wordt op (10.4.1) met een aanpassing aan de sterk negatieve eigenwaarde van de Jacobiaan.

Opmerking (IV15) Onze theorie berust essentieel op de R-consistentie en R-stabiliteit, en dus op de decompositie van de oplossing van een differentiaalvergelijking in reguliere en "sneldalende" componenten. Deze laatsten hebben we niet behandeld. Een foutenanalyse voor "sneldalende" componenten (die wel weer niet-heuristisch gedefinieerd kunnen worden, zie mijn proefschrift) is daarom zo lastig, omdat voor $|z\lambda|$ groot, λ de sterk negatieve eigenwaarde van de Jacobiaan van het stelsel, tot nu toe geen geschikt consistentie-begrip gevonden is voor deze componenten.

Zeer globaal sprekend, kan men aantonen dat deze "sneldalende" componenten weinig problemen geven als ze vrijwel uitgedempt zijn (dit geldt niet voor de trapeziumregel: men rekene daartoe het voorbeeld uit de inleiding (10.1) na met $1000 e^{-t}$ i.p.v. e^{-t}). Soms ook kan bewezen worden dat de aanwezigheid van een "grote sneldalende component" in $t=0$ toch geen "kleine" stapgrootte vereist, ook al is de "sneldalende" component nog niet uitgedempt ten opzichte van de tolerantie. Dit hangt sterk af van de methode en de differentiaalvergelijking. Enkele eenvoudige voorbeelden komen voor in mijn proefschrift.

Literatuur

In de eerste plaats zij verwezen naar mijn proefschrift. Het bepalen van oplossingen van differentiaalvergelijkingen (of recursie) die een bepaald groeikarakter hebben, is al te vinden in Ch. 13 van E.A. Coddington, N. Levinson, Theory of Ordinary Differential Equations, McGraw-Hill, New York, [1955]. De convergentietheorie voor reguliere oplossingen is een generalisatie van

de bekende theorie van Dahlquist. Zie hiervoor:

P. Henrici, Discrete Variable methods in ordinary differential equations, John Wiley and Sons, New York [1962].

Tenslotte zij verwezen naar de voorgaande bijdragen tot dit colloquium, waarin begrippen als A-stabiliteit, exponentieel fitten (deel I, §2), de formules van Liniger en Willoughby (deel I, (2,13)), Taylor-methode en semi-impliciete Runge-Kutta methoden (deel I §3) al behandeld zijn.

11. Testresultaten van methoden voor stijve vergelijkingen

Inleiding

In deze laatste voordracht geven we de resultaten van een vergelijkend onderzoek van enkele methoden voor het oplossen van stijve differentiaal-vergelijkingen die op dit colloquium ter sprake zijn gekomen.

Op het Mathematisch Centrum zijn van deze methoden de volgende ALGOL 60 versies beschikbaar:

- 1 **MULTISTEP;**
 Een lineaire meerstapsmethode gebaseerd op de formules van Adams-Moulton en Curtiss-Hirschfelder;
 Referenties: Hoofdstuk 4 van dit colloquiumverslag en [1];
- 2 **EFSIRK;**
 Een semi-impliciete, exponentieel aangepaste Runge-Kutta methode;
 Referentie: Hoofdstuk 8 van dit colloquiumverslag;
- 3 **Exponentially fitted runge kutta;**
 Een expliciete, exponentieel aangepaste Runge-Kutta methode;
 Referenties: Hoofdstuk 7 van het colloquiumverslag en [2];
- 4 **EFERK;**
 Een expliciete, exponentieel aangepaste Runge-Kutta methode die gebruik maakt van de jacobiaan;
 Referentie: Hoofdstuk 8 van dit colloquiumverslag;
- 5 **Liniger 1;**
 Een impliciete methode van orde 1 met exponentiële aanpassing in één punt;
 Referenties: Hoofdstuk 7 van dit colloquiumverslag en [3];
- 6 **Liniger 2;**
 Een impliciete methode van orde 2 met exponentiële aanpassing in twee punten;
 Referenties: Dezelfde als voor Liniger 1.

De bij de testvoorbeelden opgenomen methode liniger 3 is een modificatie van liniger 2 die ontstaat door één van de fit-punten gelijk 0 te kiezen.

De methoden `expfrk <i><j>`, die bij de testvoorbeelden worden gebruikt,

zijn versies van exponentially fitted runge kutta waarbij i de orde aan-
geeft en j de graad van de exponentiële aanpassing.

Van de bovenstaande methoden geven we in Tabel 11.1 de belangrijkste
stuurparameters.

Tabel 11.1

Belangrijkste stuurparameters van de geteste methoden

Methode	delta	phi	delta 2	lin	eval	itmax	eps
MULTISTEP						*	*
EFSIRK	*			*			
EXPFRK<i><j>	*	*					
EFERK	*	*		*			
LINIGER 1	*				*	*	*
LINIGER 2	*		*		*	*	*
LINIGER 3	*				*	*	*

We geven nu een korte beschrijving van de parameters in Tabel 11.1.

delta, phi, delta 2

Deze parameters bepalen de punten waarin exponentieel wordt aangepast. In
de testvoorbeelden hebben we steeds gefit in de in modulus grootste eigen-
waarde (met negatief reëel deel) van de jacobiaan.

lin, eval

De parameter lin geeft aan of het stelsel van differentiaalvergelijkingen
lineair is. Als dit het geval is, wordt slechts eenmaal de jacobiaan geëva-
lueerd.

eval is een boolean procedure met als parameter het aantal, per stap, uit-
gevoerde Newton-iteraties. Bij niet lineaire problemen heeft eval het vol-
gende effect: per stap wordt in liniger 1 en liniger 2 één jacobiaan geëva-
lueerd, terwijl in liniger 3 dit iedere iteratieslag gebeurt.

itmax, eps

itmax geeft het maximale aantal iteratieslagen in één stap, waarbij eps de
gewenste nauwkeurigheid van het iteratieproces aangeeft. In alle voorbeel-

den kozen we $\text{eps} = 10^{-4}$; itmax kreeg bij lineaire stelsels de waarde 1 en voor niet-lineaire stelsels de waarde 3.

Stapkeuze

Hoewel sommige van de genoemde procedures zijn uitgerust met een automatisch stapkeuze-mechanisme, hebben we (om een vergelijking tussen alle methoden mogelijk te maken) gekozen voor de volgende stap-strategie.

Allereerst splitsten we het integratie-interval in twee stukken: het begin-interval (x_0, x_1) en het eind-interval (x_1, x_2) . Het begin-interval, waarop het inschakelverschijnsel grotendeels het gedrag van de oplossing bepaalt, integreerden we met een kleine constante stap h_1 . Op het eind-interval, waarop de invloed van het inschakelverschijnsel een verwaarloosbare rol speelt, integreerden we met een stap h_2 ($h_2 > h_1$).

In de volgende voorbeelden geven we eerst de differentiaalvergelijking met beginvoorwaarden en de referentie-oplossing.

Vervolgens specificeren we de waarden van de voornaamste stuurparameters, waarna, in 7 tabellen, de testresultaten volgen.

De eerste 3 tabellen geven informatie over de integratie van (x_0, x_1) , terwijl de volgende 4 tabellen resultaten geven over (x_0, x_2) .

In kolom 1 tot en met 9 van iedere tabel staan respectievelijk de methode, het aantal juiste cijfers in component i ($\text{sd} < i >$), het aantal uitgevoerde functie-evaluaties (fev), het aantal jacobiaan-evaluaties (jev), het aantal LU ontbindingen (LU) en de gemeten rekentijd.

De niet-optimale programmering en swap-tijden (de methoden zijn geprogrammeerd en getest in ALGOL 60 op de EL X8 van het Mathematisch Centrum) doen echter, in belangrijke mate, afbreuk aan de betrouwbaarheid van de gegeven reketijden.

Het aantal goede cijfers in component i berekenen we als volgt

$$\text{sd} < i > = -10 \log \left| \frac{y_i}{y_{\text{ref}_i}} - 1 \right|, \quad |\text{sd} < i >| \leq 12.$$

Bijna alle testvoorbeelden zijn ontleend aan een memorandum [7-3-1972] opgesteld door T.E. Hull, W.H. Enright en A.E. Sedgewick, met richtlijnen en voorbeelden voor het testen van procedures voor het oplossen van beginwaarde-problemen. Hieruit zijn ook verschillende suggesties voor de te volgen procedure bij het testen overgenomen.

De eerste 6 voorbeelden zijn lineaire differentiaalvergelijkingen waarvan de laatste 2 stelsels zijn met een orde van 10 respectievelijk 51.

De volgende 4 voorbeelden zijn niet-lineaire stelsels, waaronder een van de orde 12.

Als laatste voorbeeld kozen we een eenvoudige niet-stijve differentiaalvergelijking.

Opmerking:

Zoals hierboven is vermeld, werd over elk opgegeven integratie-interval het integratieproces met een constante staplengte uitgevoerd. De stapkeuze- en ordekeuze-strategie van de procedure MULTISTEP bleken zodanig met elkaar verweven, dat, door het opleggen van een bepaalde staplengte, ook de orde van de integratieformule gefixeerd was.

Zo werd van de Curtiss-Hirschfelder formules alleen de eerste orde formule gebruikt.

In feite werd dus niet een echte multistepmethode gebruikt, maar de backward-Euler methode.

In de testresultaten geven we het gebruik van deze methode dan ook aan met bw. euler.

Voorbeeld 1

Differentiaalvergelijking:

$$y' = \begin{bmatrix} -1/2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -100 & 0 \\ 0 & 0 & 0 & -90 \end{bmatrix} y .$$

Beginvoorwaarden:

$$y(0) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} .$$

Integratie-interval: [0,20].

Referentie-oplossing:

$$y = \begin{bmatrix} e^{-\frac{x}{2}} \\ e^{-x} \\ e^{-100x} \\ e^{-90x} \end{bmatrix} .$$

Stuurparameters: delta = 100, phi = π , delta 2 = 90.

results for integration until $x = +.25_{10}-0$

$h1 = +.31250_{10}-1$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+3.0	+2.4	-5.9	-5.1	9	1	1	.5
efsirk	+8.7	+7.5	+10.7	+6	16	1	1	.5
expfrk23	+6.4	+5.5	+9.0	+1.6	40	0	0	.3
expfrk22	+6.0	+5.1	+8.6	+1	32	0	0	.3
expfrk13	+3.9	+3.3	+8.9	+1.4	32	0	0	.3
expfrk11	+3.3	+2.7	+9.7	+3	16	0	0	.2
eferk	+8.5	+7.3	+8.8	+5	16	1	0	.5
liniger1	+3.4	+2.8	+10.1	-7	8	1	1	.4
liniger2	+6.5	+5.6	+10.0	+9.8	8	1	1	.5
liniger3	+8.7	+7.5	+10.2	+6	8	1	1	.5

$h1 = +.15625_{10}-1$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+3.3	+2.7	-4.3	-3.7	17	1	1	.9
efsirk	+9.9	+8.7	+11.1	+1.8	32	1	1	.8
expfrk23	+7.7	+6.8	+11.0	+3.4	80	0	0	.6
expfrk22	+7.0	+6.1	+9.5	+1.7	64	0	0	.5
expfrk13	+4.8	+4.2	+10.2	+2.8	64	0	0	.5
expfrk11	+3.7	+3.1	+10.5	+1	32	0	0	.3
eferk	+9.9	+8.7	+10.8	+2.3	32	1	0	1.0
liniger1	+3.9	+3.3	+10.2	+3	16	1	1	.7
liniger2	+7.6	+6.7	+10.1	+11.2	16	1	1	.9
liniger3	+9.9	+8.7	+10.3	+1.8	16	1	1	.9

$h1 = +.78125_{10}-2$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+3.6	+3.0	-2.8	-2.4	33	1	1	1.7
efsirk	+11.1	+9.9	+10.9	+3.1	64	1	1	1.4
expfrk23	+12.0	+12.0	+1.7	+2.0	160	0	0	1.0
expfrk22	+12.0	+11.1	+8	+1.1	128	0	0	.8
expfrk13	+12.0	+11.1	+8	+1.1	128	0	0	.8
expfrk11	+6.5	+5.6	-1.9	-1.3	64	0	0	.6
eferk	+12.0	+12.0	+1.7	+2.0	64	1	0	1.8
liniger1	+4.5	+3.9	+10.1	+1.0	32	1	1	1.2
liniger2	+11.2	+9.9	+10.9	+3.1	32	1	1	1.7
liniger3	+11.2	+9.9	+10.9	+3.1	32	1	1	1.7

results for integration until $x = +.20_{10}+2$

$h1 = +.31250_{10}-1$ $h2 = +.12500_{10}-0$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+5	-3	-12.0	-12.0	167	2	2	8.1
efsirk	+4.6	+3.4	-1.0	-12.0	332	1	2	6.8
expfrk23	+2.5	+1.6	-.7	-12.0	830	0	0	5.0
expfrk22	+2.4	+1.5	-.6	-12.0	664	0	0	4.2
expfrk13	+.8	+.3	-1.2	-12.0	664	0	0	4.2
expfrk11	+.6	+.2	-.5	-12.0	332	0	0	2.7
eferk	+4.2	+3.0	-.5	-12.0	332	1	0	10.3
liniger1	+.5	-.2	+.0	-12.0	166	1	2	5.9
liniger2	+2.5	+1.6	+.0	+.0	166	1	2	8.5
liniger3	+4.6	+3.4	+.0	-12.0	166	1	2	8.6

$h1 = +.15625_{10}-1$ $h2 = +.12500_{10}-0$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+5	-3	-12.0	-12.0	175	2	2	8.9
efsirk	+4.6	+3.4	-.8	-12.0	348	1	2	7.4
expfrk23	+2.5	+1.6	-2.0	-12.0	870	0	0	5.4
expfrk22	+2.4	+1.5	-1.6	-12.0	696	0	0	4.6
expfrk13	+.8	+.3	-.6	-12.0	696	0	0	4.6
expfrk11	+.6	+.2	-.7	-12.0	348	0	0	2.9
eferk	+4.2	+3.0	-.3	-12.0	348	1	0	10.6
liniger1	+.5	-.2	+.0	-12.0	174	1	2	5.9
liniger2	+2.5	+1.6	+.0	+.0	174	1	2	8.6
liniger3	+4.6	+3.4	+.0	-12.0	174	1	2	8.6

$h1 = +.15625_{10}-1$ $h2 = +.62500_{10}-1$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+.8	+.1	-12.0	-12.0	333	2	2	15.9
efsirk	+5.7	+4.5	+.0	+.0	664	1	2	13.4
expfrk23	+3.4	+2.5	-.6	-.5	1660	0	0	9.8
expfrk22	+3.2	+2.3	-1.0	-12.0	1328	0	0	8.3
expfrk13	+1.3	+.7	-.8	-.3	1328	0	0	8.4
expfrk11	+1.0	+.4	-.3	-12.0	664	0	0	5.3
eferk	+5.4	+4.2	+12.0	-12.0	664	1	0	19.7
liniger1	+1.0	+.3	+.0	-12.0	332	1	2	11.1
liniger2	+3.4	+2.5	+.0	+.0	332	1	2	16.2
liniger3	+5.7	+4.5	+.0	+.0	332	1	2	16.1

$$h1 = +.78125_{10} - 2 \quad h2 = +.62500_{10} - 1$$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+.8	+.1	-12.0	-12.0	349	2	2	16.6
efsirk	+5.7	+4.5	+.0	+.0	696	1	2	14.0
expfrk23	+3.4	+2.5	-1.0	+.0	1740	0	0	10.3
expfrk22	+3.2	+2.3	+.0	-12.0	1392	0	0	8.7
expfrk13	+1.3	+.7	+12.0	+12.0	1392	0	0	8.7
expfrk11	+1.0	+.4	+12.0	-12.0	696	0	0	5.6
eferk	+5.4	+4.2	-.5	-12.0	696	1	0	20.5
liniger1	+1.0	+.3	+.0	-12.0	348	1	2	12.0
liniger2	+3.4	+2.5	+.0	+.0	348	1	2	17.0
liniger3	+5.7	+4.5	+.0	+.0	348	1	2	17.4

Opmerkingen:

Op het begininterval neemt bij staphalvering de nauwkeurigheid in het algemeen toe.

De vierde component wordt slechts goed gerepresenteerd door Liniger 2 omdat bij deze methode ook in het punt -90 wordt aangepast.

Bij de kleinste staplengte gaat deze methode over in Liniger 3 omdat, in gevallen waarbij het produkt $h1 * \delta^2$ kleiner dan 1 is, slechts in een punt wordt aangepast.

Enkele resultaten berusten min of meer op toeval. We doelen hierbij op de 12 juiste decimalen die sommige procedures voor de 3e of 4e component in $xe2$ geven. Ook bij enkele volgende voorbeelden zal dit soms het geval zijn.

Voorbeeld 2

Differentiaalvergelijking:

$$y' = \begin{bmatrix} -1 & 10 & 0 & 0 \\ -10 & -1 & 0 & 0 \\ 0 & 0 & -100 & 100 \\ 0 & 0 & -100 & -100 \end{bmatrix} y .$$

Beginvoorwaarden:

$$y(0) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} .$$

Integratie-interval: $[0, 20]$.

Referentie-oplossing:

$$y_{1,2} = \frac{\cos 10x \pm \sin 10x}{e^x} ,$$

$$y_{3,4} = \frac{\cos 100x \pm \sin 100x}{e^{100x}} .$$

Stuurparameters: $\delta = 100\sqrt{2}$, $\phi = \frac{3}{4}\pi$, $\delta_2 = -\frac{3}{4}\pi$.

results for integration until $x = +.25_{10}-0$

$h_1 = +.78125_{10}-2$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+6	+1.1	-1.8	-.7	33	1	1	1.8
efsirk	+5.1	+5.8	+1.1	+1.8	64	1	1	1.5
expfrk24	+4.8	+5.8	+9.5	+11.1	192	0	0	1.4
expfrk22	+3.1	+4.3	+9.4	+10.6	128	0	0	1.0
expfrk14	+3.1	+3.6	+10.4	+9.9	160	0	0	1.2
expfrk12	+1.7	+2.2	+10.0	+10.0	96	0	0	.8
eferk	+5.2	+5.9	+9.9	+9.9	64	1	0	1.9
liniger1	+1.3	+1.8	-.9	-1.0	32	1	1	1.2
liniger2	+3.8	+4.9	+8.8	+9.5	32	1	1	1.8
liniger3	+5.1	+5.8	+1.1	+1.8	32	1	1	1.8

$h_1 = +.39062_{10}-2$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+9	+1.3	-.7	-.6	65	1	1	3.4
efsirk	+6.3	+7.0	+2.3	+3.0	128	1	1	2.9
expfrk24	+10.5	+11.4	+3.4	+3.8	384	0	0	2.5
expfrk22	+6.5	+7.5	+1.6	+1.4	256	0	0	1.8
expfrk14	+8.8	+9.6	+2.3	+2.9	320	0	0	2.1
expfrk12	+4.7	+5.3	+.7	+.3	192	0	0	1.4
eferk	+8.8	+9.6	+2.3	+2.9	128	1	0	3.6
liniger1	+1.9	+2.4	+.9	-.1	64	1	1	2.3
liniger2	+5.0	+6.1	+8.5	+9.4	64	1	1	3.3
liniger3	+6.3	+7.0	+2.3	+3.0	64	1	1	3.3

$h_1 = +.19531_{10}-2$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+1.2	+1.6	+.7	-.4	129	1	1	6.4
efsirk	+7.5	+8.2	+3.5	+4.2	256	1	1	5.4
expfrk24	+10.6	+11.4	+5.2	+5.9	768	0	0	4.7
expfrk22	+7.8	+8.7	+3.0	+2.6	512	0	0	3.4
expfrk14	+10.2	+11.0	+3.9	+4.3	640	0	0	4.1
expfrk12	+5.6	+6.2	+1.5	+1.4	384	0	0	2.8
eferk	+10.2	+11.1	+3.9	+4.3	256	1	0	7.1
liniger1	+2.5	+3.0	+1.2	+.6	128	1	1	4.4
liniger2	+6.2	+7.3	+7.4	+8.4	128	1	1	6.4
liniger3	+7.5	+8.2	+3.5	+4.2	128	1	1	6.4

results for integration until $x = +.20_{10}+2$

$h1 = +.78125_{10}-2$ $h2 = +.62500_{10}-1$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+0	+0	-12.0	-12.0	349	2	2	16.8
efsirk	+6	+5	-12.0	-12.0	696	1	2	15.9
expfrk24	-3	-2	-1.6	-1.8	2088	0	0	12.9
expfrk22	-1.0	+2	-1.3	-6	1392	0	0	9.3
expfrk14	-12.0	-12.0	-3	-1.0	1740	0	0	11.1
expfrk12	-12.0	-12.0	-7	-7	1044	0	0	7.6
eferk	+2	+2	+12.0	+0	696	1	0	18.9
liniger1	+0	+0	-12.0	-12.0	348	1	2	11.8
liniger2	+0	+0	+0	+0	348	1	2	18.2
liniger3	+6	+5	-12.0	-12.0	348	1	2	17.3

$h1 = +.39062_{10}-2$ $h2 = +.62500_{10}-1$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+0	+0	-12.0	-12.0	381	2	2	19.3
efsirk	+6	+5	-12.0	-12.0	760	1	2	15.9
expfrk24	-3	-2	-1.9	-2.0	2280	0	0	14.9
expfrk22	-1.0	+2	-1.1	-1.4	1520	0	0	10.1
expfrk14	-12.0	-12.0	-1.3	-1.5	1900	0	0	13.0
expfrk12	-12.0	-12.0	-3	-1.2	1140	0	0	8.3
eferk	+2	+2	-3	+12.0	760	1	0	20.9
liniger1	+0	+0	-12.0	-12.0	380	1	2	12.9
liniger2	+0	+0	+0	+0	380	1	2	18.7
liniger3	+6	+5	-12.0	-12.0	380	1	2	18.8

$h1 = +.39062_{10}-2$ $h2 = +.31250_{10}-1$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	-0	+0	-12.0	-12.0	697	2	2	33.3
efsirk	+1.6	+1.5	-5	-5	1392	1	2	28.6
expfrk24	-0	+1.5	-8	-6	4176	0	0	26.5
expfrk22	-5	+4	-5	-5	2784	0	0	19.5
expfrk14	-2.4	-2.1	-6	-5	3480	0	0	22.1
expfrk12	-6.7	-6.2	-6	-3	2088	0	0	15.2
eferk	+1.9	+1.2	+12.0	+0	1392	1	0	37.7
liniger1	+0	+0	-12.0	-12.0	696	1	2	23.5
liniger2	-2	+6	+0	+0	696	1	2	34.1
liniger3	+1.6	+1.5	+0	+0	696	1	2	34.3

$$h1 = +.19531_{10} - 2 \quad h2 = +.31250_{10} - 1$$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	-.0	+.0	-12.0	-12.0	761	2	2	36.4
efsirk	+1.6	+1.5	+12.0	+.0	1520	1	2	31.2
expfrk24	-.0	+1.5	-.3	-1.0	4560	0	0	27.8
expfrk22	-.5	+.4	-.5	+.0	3040	0	0	20.7
expfrk14	-2.4	-2.1	-.3	-.6	3800	0	0	26.7
expfrk12	-6.7	-6.2	-.5	+.0	2280	0	0	17.7
eferk	+1.9	+1.2	+.0	+12.0	1520	1	0	43.3
liniger1	+.0	+.0	-12.0	-12.0	760	1	2	25.6
liniger2	-.2	+.6	+.0	+.0	760	1	2	40.3
liniger3	+1.6	+1.5	+.0	+.0	760	1	2	38.5

Opmerkingen:

Hetzelfde verschijnsel, waarvan in voorbeeld 1 bij Liniger 2 sprake was, doet zich nu voor bij de methoden expfrk<i><j> en eferk.

Voor grote $h1$ representeren deze formules de laatste componenten goed omdat ze aanpassing in complexe punten toelaten hetgeen ook geldt voor Liniger 2. Uit de gegevens over het eindinterval blijkt dat daarop met een zeer kleine stap moet worden geïntegreerd om nog enige nauwkeurigheid in de oplossing te verkrijgen.

Deze moeilijkheid wordt veroorzaakt door de bijna zuiver imaginaire eigenwaarden $-1 \pm 10i$.

Voorbeeld 3

Differentiaalvergelijking:

$$y' = \begin{bmatrix} -10^5 & +10^2 & -10 & 1 \\ 0 & -10^4 & +10 & -10 \\ 0 & 0 & -10 & 10 \\ 0 & 0 & 0 & -1 \end{bmatrix} y .$$

Beginvoorwaarden:

$$y(0) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} .$$

Integratie-interval: $[0, 20]$.

Referentie-oplossing:

$$y_1 = c_1 e^{-10^5 x} + c_2 e^{-10^4 x} + c_3 e^{-10x} + c_4 e^{-x},$$

$$y_2 = 900c_2 e^{-10^4 x} - \frac{1}{9 \times 999} e^{-10x} + \frac{10}{9 \times 9999} e^{-x},$$

$$y_3 = (10e^{-x} - e^{-10x})/9,$$

$$y_4 = e^{-x},$$

met:

$$c_2 = (1 + \frac{1}{12208779})/900, \quad c_3 = (\frac{1}{9} - \frac{10}{9 \times 999})/9999,$$

$$c_4 = (\frac{1000}{9 \times 9999} - \frac{100}{9} + 1)/99999, \quad c_1 = 1 - c_2 - c_3 - c_4.$$

Stuurparameters: $\delta = 10^5$, $\phi = \pi$, $\delta_2 = 10^4$.

results for integration until $x = +.25_{10^{-3}}$

$h1 = +.78125_{10^{-5}}$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	-.9	+1.0	+8.0	+9.0	33	1	1	1.7
efsirk	+4.1	+5.9	+11.7	+11.7	64	1	2	1.6
expfrk23	+6.0	+8.0	+12.0	+11.7	160	0	0	1.1
expfrk22	+4.2	+6.1	+12.0	+11.7	128	0	0	.9
expfrk13	+4.2	+6.1	+12.0	+11.7	128	0	0	.9
expfrk11	+7	+2.6	+12.0	+11.7	64	0	0	.6
eferk	+6.0	+8.0	+11.7	+11.7	64	1	0	1.9
liniger1	+1	+1.9	+8.9	+9.9	32	1	2	1.3
liniger2	+4.1	+5.9	+11.3	+11.6	32	1	2	1.8
liniger3	+4.1	+5.9	+11.3	+11.6	32	1	2	1.8

$h1 = +.39062_{10^{-5}}$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	-.6	+1.3	+8.3	+9.3	65	1	1	3.4
efsirk	+5.3	+7.1	+11.4	+12.0	130	1	2	3.0
expfrk23	+7.5	+9.5	+12.0	+11.7	320	0	0	2.2
expfrk22	+5.4	+7.3	+12.0	+11.7	256	0	0	1.8
expfrk13	+5.4	+7.3	+12.0	+11.7	256	0	0	1.8
expfrk11	+1.3	+3.2	+12.0	+11.7	128	0	0	1.2
eferk	+7.5	+9.5	+11.4	+12.0	128	1	0	4.6
liniger1	+7	+2.5	+9.5	+10.5	64	1	2	2.4
liniger2	+5.3	+7.1	+10.6	+10.8	64	1	2	3.4
liniger3	+5.3	+7.1	+10.6	+10.8	64	1	2	3.4

$h1 = +.19531_{10^{-5}}$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	-.3	+1.6	+8.6	+9.7	129	1	1	6.3
efsirk	+6.5	+8.3	+12.0	+12.0	258	1	2	5.7
expfrk23	+8.4	+10.3	+12.0	+11.7	645	0	0	4.4
expfrk22	+6.6	+8.5	+12.0	+11.7	516	0	0	3.7
expfrk13	+6.6	+8.5	+12.0	+11.7	516	0	0	3.7
expfrk11	+1.9	+3.8	+12.0	+11.7	258	0	0	3.1
eferk	+8.4	+10.3	+12.0	+12.0	256	1	0	7.2
liniger1	+1.3	+3.1	+10.1	+11.0	128	1	2	4.5
liniger2	+6.5	+8.3	+10.3	+10.4	128	1	2	6.5
liniger3	+6.5	+8.3	+10.3	+10.4	128	1	2	6.5

results for integration until $x = +.20_{10}+2$

$h1 = +.78125_{10}-5$ $h2 = +.79999_{10}-1$

methode	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	-.1	-.1	-.1	-.1	283	2	2	13.7
efsirk	+3.9	+3.9	+3.9	+3.9	566	1	3	11.8
expfrk23	-12.0	-12.0	+1.6	+1.6	1410	0	0	8.9
expfrk22	-12.0	-12.0	+1.6	+1.6	1128	0	0	7.5
expfrk13	-12.0	-12.0	+.2	+.2	1128	0	0	7.6
expfrk11	-12.0	-12.0	+.2	+.2	564	0	0	4.8
eferk	-12.0	-12.0	+3.3	+3.3	564	1	0	17.0
liniger1	-.1	-.1	-.1	-.1	282	1	3	9.9
liniger2	+1.7	+1.7	+1.7	+1.7	282	1	3	15.1
liniger3	+3.9	+3.9	+3.9	+3.9	282	1	3	15.4

$h1 = +.39062_{10}-5$ $h2 = +.79999_{10}-1$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	-.1	-.1	-.1	-.1	315	2	2	15.5
efsirk	+3.9	+3.9	+3.9	+3.9	630	1	3	14.7
expfrk23	-12.0	-12.0	+1.6	+1.6	1575	0	0	10.6
expfrk22	-12.0	-12.0	+1.6	+1.6	1260	0	0	8.4
expfrk13	-12.0	-12.0	+.2	+.2	1260	0	0	8.4
expfrk11	-12.0	-12.0	+.2	+.2	630	0	0	5.4
eferk	-12.0	-12.0	+3.3	+3.3	630	1	0	18.9
liniger1	-.1	-.1	-.1	-.1	315	1	3	10.8
liniger2	+1.7	+1.7	+1.7	+1.7	315	1	3	15.7
liniger3	+3.9	+3.9	+3.9	+3.9	315	1	3	15.7

$h1 = +.39062_{10}-5$ $h2 = +.40000_{10}-1$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+.3	+.3	+.3	+.3	565	2	2	27.1
efsirk	+4.8	+4.8	+4.8	+4.8	1130	1	3	23.3
expfrk23	-12.0	-12.0	+2.3	+2.3	2825	0	0	17.8
expfrk22	-12.0	-12.0	+2.3	+2.3	2260	0	0	16.2
expfrk13	-12.0	-12.0	+.5	+.5	2260	0	0	15.3
expfrk11	-12.0	-12.0	+.5	+.5	1130	0	0	10.4
eferk	-12.0	-12.0	+4.3	+4.3	1130	1	0	36.1
liniger1	+.3	+.3	+.3	+.3	565	1	3	20.2
liniger2	+2.3	+2.3	+2.3	+2.3	565	1	3	27.9
liniger3	+4.8	+4.8	+4.8	+4.8	565	1	3	27.9

$$h1 = +.19531 \cdot 10^{-5} \quad h2 = +.40000 \cdot 10^{-1}$$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+3	+3	+3	+3	629	2	2	30.1
efsirk	+4.8	+4.8	+4.8	+4.8	1258	1	3	25.9
expfrk23	-12.0	-12.0	+2.3	+2.3	3145	0	0	19.7
expfrk22	-12.0	-12.0	+2.3	+2.3	2516	0	0	16.7
expfrk13	-12.0	-12.0	+5	+5	2516	0	0	17.6
expfrk11	-12.0	-12.0	+5	+5	1258	0	0	10.6
eferk	-12.0	-12.0	+4.3	+4.3	1258	1	0	39.9
liniger1	+3	+3	+3	+3	629	1	3	22.7
liniger2	+2.3	+2.3	+2.3	+2.3	629	1	3	33.4
liniger3	+4.8	+4.8	+4.8	+4.8	629	1	3	31.0

Opmerkingen:

De slechte resultaten (op het eindinterval) van de formules $\text{expfrk}_{i < j}$ en eferk worden veroorzaakt door instabiliteit bij de gebruikte staplengte (zie (7.24)). Uit de analytische oplossing zien we dat voor grote x alle componenten voorgesteld worden door $c_i e^{-x}$ ($i=1,2,3,4$). Dit is in de resultaten gemakkelijk terug te vinden.

Voorbeeld 4

Differentiaalvergelijking:

$$y' = \begin{bmatrix} -10^4 & 10^3 & 0 & 0 \\ -10^3 & -10^4 & 0 & 0 \\ 0 & 0 & -10 & 100 \\ 0 & 0 & -100 & -10 \end{bmatrix} y .$$

Beginvoorwaarden:

$$y(0) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} .$$

Integratie-interval: $[0, 20]$.

Referentie-oplossing:

$$y_{1,2} = \frac{\cos 10^3 x \pm \sin 10^3 x}{e^{10^4 x}} ,$$

$$y_{3,4} = \frac{\cos 100x \pm \sin 100x}{e^{10x}} .$$

Stuurparameters: $\delta = 1000\sqrt{101}$, $\phi = \pi - \arctan(.1)$, $\delta_2 = -\phi$.

results for integration until $x = +.25_{10}-2$

$h1 = +.78125_{10}-4$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	-3.6	-2.6	+3.0	+3.2	33	1	1	1.7
efsirk	+2.0	+3.4	+9.8	+10.5	64	1	2	2.9
expfrk24	+2.0	+2.7	+11.7	+11.4	192	0	0	1.3
expfrk22	+3	+8	+11.7	+10.7	128	0	0	1.0
expfrk14	+1.2	+1.7	+11.7	+11.4	160	0	0	1.2
expfrk12	+2	+2	+8.2	+8.8	96	0	0	.8
eferk	+1.2	+1.7	+12.0	+11.1	64	1	0	1.9
liniger1	+0	+2.1	+3.9	+4.1	32	1	2	1.3
liniger2	+9.1	+9.2	+8.5	+7.6	32	1	2	1.9
liniger3	+2.0	+3.4	+9.8	+10.6	32	1	2	1.8

$h1 = +.39063_{10}-4$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	-2.3	-1.6	+3.3	+3.5	65	1	1	3.3
efsirk	+3.3	+4.6	+11.1	+11.7	128	1	2	3.0
expfrk24	+4.0	+4.7	+11.7	+11.1	384	0	0	2.6
expfrk22	+1.7	+2.2	+11.7	+11.0	256	0	0	1.9
expfrk14	+2.8	+3.4	+11.7	+11.1	320	0	0	2.2
expfrk12	+8	+1.1	+9.1	+9.8	192	0	0	1.5
eferk	+2.8	+3.4	+12.0	+11.3	128	1	0	3.7
liniger1	+6	+2.5	+4.5	+4.7	64	1	2	3.2
liniger2	+8.6	+8.7	+9.7	+8.8	64	1	2	3.4
liniger3	+3.3	+4.6	+11.1	+11.1	64	1	2	3.4

$h1 = +.19531_{10}-4$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	-1.2	-.9	+3.6	+3.8	129	1	1	6.4
efsirk	+4.5	+5.8	+12.0	+10.9	256	1	2	5.5
expfrk24	+5.9	+6.5	+11.3	+10.9	768	0	0	4.9
expfrk22	+3.0	+3.5	+11.3	+10.9	512	0	0	3.6
expfrk14	+4.4	+5.0	+11.3	+10.9	640	0	0	5.1
expfrk12	+1.8	+2.0	+10.0	+11.0	384	0	0	3.0
eferk	+4.4	+5.0	+12.0	+10.9	256	1	0	7.4
liniger1	+1.3	+3.1	+5.1	+5.3	128	1	2	4.5
liniger2	+7.6	+7.7	+10.7	+10.1	128	1	2	6.6
liniger3	+4.5	+5.8	+11.6	+10.8	128	1	2	7.4

results for integration until $x = +.20_{10}+2$

$h1 = +.78125_{10}-4$ $h2 = +.79990_{10}-1$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+0	+0	+0	+0	283	2	2	13.8
efsirk	-2.1	-2.3	+0	+0	566	1	3	12.1
expfrk24	-5.6	-4.7	-12.0	-12.0	1698	0	0	11.8
expfrk22	-5.0	-4.8	-12.0	-12.0	1132	0	0	7.9
expfrk14	-2.2	-1.8	-12.0	-12.0	1415	0	0	9.4
expfrk12	-2.2	-2.8	-12.0	-12.0	849	0	0	7.3
eferk	-2.2	-2.2	-12.0	-12.0	564	1	0	15.8
liniger1	+0	+0	+0	+0	282	1	3	9.8
liniger2	+12.0	+0	+0	+0	282	1	3	15.3
liniger3	-.3	+0	+0	+0	282	1	3	14.3

$h1 = +.39063_{10}-4$ $h2 = +.79990_{10}-1$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+12.0	+0	+0	+0	315	2	2	16.2
efsirk	-2.6	-2.8	+0	+0	630	1	3	13.3
expfrk24	-5.7	-5.4	-12.0	-12.0	1890	0	0	13.0
expfrk22	-4.5	-4.5	-12.0	-12.0	1260	0	0	8.8
expfrk14	-3.0	-2.8	-12.0	-12.0	1575	0	0	10.4
expfrk12	-2.0	-2.0	-12.0	-12.0	945	0	0	8.1
eferk	-2.6	-2.6	-12.0	-12.0	628	1	0	17.7
liniger1	+0	+0	+0	+0	314	1	3	11.8
liniger2	+12.0	+0	+0	+0	314	1	3	15.8
liniger3	-.3	+0	+0	+0	314	1	3	16.7

$h1 = +.39063_{10}-4$ $h2 = +.39995_{10}-1$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+0	-.3	+0	+0	565	2	2	28.1
efsirk	-2.4	-2.4	+0	+0	1130	1	3	23.8
expfrk24	-4.8	-4.4	-12.0	-12.0	3390	0	0	23.0
expfrk22	-5.0	-4.8	-12.0	-12.0	2260	0	0	16.9
expfrk14	-2.0	-2.6	-12.0	-12.0	2825	0	0	18.6
expfrk12	-1.6	-2.0	-12.0	-12.0	1695	0	0	13.6
eferk	-2.2	-2.6	-12.0	-12.0	1128	1	0	32.2
liniger1	+0	+0	+0	+0	564	1	3	19.3
liniger2	+0	+0	+0	+0	564	1	3	28.5
liniger3	+12.0	+12.0	+0	+0	564	1	3	28.9

$$h1 = +.19531_{10^{-4}} \quad h2 = +.39995_{10^{-1}}$$

method	sd1	sd2	sd3	sd4	fev	jev	lu	time
bw.euler	+0	+0	+0	+0	629	2	2	31.0
efsirk	-2.7	-2.4	+0	+0	1258	1	3	26.3
expfrk24	-5.1	-5.5	-12.0	-12.0	3774	0	0	24.0
expfrk22	-4.9	-4.2	-12.0	-12.0	2516	0	0	17.4
expfrk14	-3.2	-3.0	-12.0	-12.0	3145	0	0	20.7
expfrk12	-2.4	-1.7	-12.0	-12.0	1887	0	0	14.1
eferk	-1.3	-2.4	-12.0	-12.0	1256	1	0	35.0
liniger1	+0	+0	+0	+0	628	1	3	21.4
liniger2	+0	+0	+0	+0	628	1	3	30.8
liniger3	+0	+0	+0	+0	628	1	3	31.0

Opmerkingen:

De resultaten vertonen veel overeenkomst met die van voorbeeld 2. Ook hier blijkt weer dat het eindinterval met zeer kleine stap moet worden geïntegreerd, indien men althans geïnteresseerd is in relatieve precisie. De absolute fout in de oplossing in het punt xe2 is namelijk van de orde 10^{-80} .

Voorbeeld 5

Differentiaalvergelijking:

$$\begin{bmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_{51} \end{bmatrix} = \begin{bmatrix} -2 & 1 & & & \\ & 1 & -2 & 1 & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 & -2 & 1 \\ & & & & & 1 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{51} \end{bmatrix} .$$

Beginvoorwaarden:

$$y(0) = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} .$$

Integratie-interval: [0,20].

Referentie-oplossing:

$$\begin{aligned}
 {}_1y(10) &= .8750622218358_{10}^{-2} , & {}_1y(20) &= .3124111453740_{10}^{-2} , \\
 {}_6y(10) &= .2155042988693_{10}^{-1} , & {}_6y(20) &= .1204549525745_{10}^{-1} , \\
 {}_{27}y(10) &= .1501604355442_{10}^{-7} , & {}_{27}y(20) &= .1162956710418_{10}^{-4} , \\
 {}_{51}y(10) &= .4325334300072_{10}^{-22} , & {}_{51}y(20) &= .1891330691035_{10}^{-13} .
 \end{aligned}$$

Stuurparameters: $\text{delta} = 2 + 2 \cos(\pi/51)$, $\text{delta } 2 = 2 + 2 \cos(\pi/26)$,
 $\text{phi} = \pi$.

results for integration until $x = +.10000_{10}+2$

$h1 = +.31250_{10}-0$

method	sd1	sd6	sd27	sd51	fev	jev	lu	time
bw.euler	+1.2	+2.9	-.6	-4.4	33	1	1	36.2
efsirk	+5.6	+6.1	+3.3	+1.0	64	1	1	49.5
expfrk22	+4.0	+4.7	+1.9	+.4	128	0	0	10.8
expfrk21	+3.3	+3.9	+1.3	+.1	96	0	0	7.5
expfrk12	+2.4	+4.0	+.7	+.0	96	0	0	7.5
expfrk11	+1.8	+3.6	+.3	+.0	64	0	0	5.7
eferk	+4.8	+5.3	+2.7	+.7	64	1	0	32.0
liniger1	+2.0	+3.8	+.2	-1.9	32	1	1	29.9
liniger2	+4.7	+5.4	+2.5	+.7	32	1	1	57.7
liniger3	+5.6	+6.1	+3.3	+1.0	32	1	1	57.7

$h1 = +.15625_{10}-0$

method	sd1	sd6	sd27	sd51	fev	jev	lu	time
bw.euler	+1.5	+3.5	-.2	-2.7	65	1	1	59.9
efsirk	+6.8	+7.3	+4.5	+2.3	128	1	1	76.2
expfrk22	+6.8	+7.2	+4.8	+2.0	256	0	0	20.0
expfrk21	+5.0	+5.5	+2.9	+1.0	192	0	0	14.9
expfrk12	+5.0	+5.5	+2.9	+1.0	192	0	0	16.4
expfrk11	+3.3	+3.9	+1.3	+.2	128	0	0	11.2
eferk	+6.8	+7.2	+4.8	+2.0	128	1	0	63.0
liniger1	+2.6	+4.3	+.9	-.5	64	1	1	48.4
liniger2	+6.8	+7.3	+4.5	+2.3	64	1	1	91.1
liniger3	+6.8	+7.3	+4.5	+2.3	64	1	1	88.4

$h1 = +.78125_{10}-1$

method	sd1	sd6	sd27	sd51	fev	jev	lu	time
bw.euler	+1.8	+4.0	+.2	-1.5	129	1	1	108.2
efsirk	+8.0	+8.5	+5.7	+4.7	256	1	1	129.8
expfrk22	+8.0	+8.5	+6.0	+3.4	512	0	0	39.4
expfrk21	+5.9	+6.4	+3.8	+1.8	384	0	0	31.6
expfrk12	+5.9	+6.4	+3.8	+1.8	384	0	0	31.2
expfrk11	+3.9	+4.5	+1.9	+.6	256	0	0	23.8
eferk	+8.0	+8.5	+6.0	+3.4	256	1	0	124.1
liniger1	+3.2	+4.9	+1.5	+.3	128	1	1	80.4
liniger2	+8.0	+8.5	+5.7	+4.7	128	1	1	156.1
liniger3	+8.0	+8.5	+5.7	+4.7	128	1	1	156.8

results for integration until $x = +.20000_{10}+2$

$h1 = +.31250_{10}-0$ $h2 = +.31250_{10}-0$

method	sd1	sd6	sd27	sd51	fev	jev	lu	time
bw.euler	+1.5	+1.9	+.6	-1.3	65	2	2	71.3
efsirk	+6.5	+7.5	+5.8	+2.8	128	1	1	78.2
expfrk22	+4.6	+5.2	+3.8	+1.6	256	0	0	18.9
expfrk21	+3.9	+4.5	+3.1	+1.0	192	0	0	16.6
expfrk12	+2.7	+3.0	+1.7	+.5	192	0	0	15.0
expfrk11	+2.1	+2.4	+1.1	+.1	128	0	0	12.7
eferk	+5.7	+6.8	+5.2	+2.2	128	1	0	62.6
liniger1	+2.2	+2.6	+1.3	-.2	64	1	1	45.9
liniger2	+5.2	+5.8	+4.4	+2.1	64	1	1	90.5
liniger3	+6.5	+7.5	+5.8	+2.8	64	1	1	90.4

$h1 = +.15625_{10}-0$ $h2 = +.31250_{10}-0$

method	sd1	sd6	sd27	sd51	fev	jev	lu	time
bw.euler	+1.7	+2.0	+.7	-1.0	97	2	2	95.5
efsirk	+6.7	+7.8	+6.1	+3.1	192	1	2	126.9
expfrk22	+4.9	+5.5	+4.1	+1.9	384	0	0	29.2
expfrk21	+4.2	+4.8	+3.4	+1.3	288	0	0	22.3
expfrk12	+3.0	+3.3	+2.0	+.7	288	0	0	23.8
expfrk11	+2.4	+2.7	+1.4	+.3	192	0	0	18.4
eferk	+6.0	+7.1	+5.5	+2.5	192	1	0	93.2
liniger1	+2.5	+2.8	+1.5	+.1	96	1	2	74.8
liniger2	+5.5	+6.1	+4.7	+2.4	96	1	2	145.3
liniger3	+6.7	+7.8	+6.1	+3.1	96	1	2	147.5

$h1 = +.15625_{10}-0$ $h2 = +.15625_{10}-0$

method	sd1	sd6	sd27	sd51	fev	jev	lu	time
bw.euler	+1.8	+2.2	+.9	-.6	129	2	2	120.0
efsirk	+7.7	+8.7	+7.0	+4.0	256	1	1	129.8
expfrk22	+8.0	+9.3	+7.4	+4.1	512	0	0	38.4
expfrk21	+5.9	+7.0	+5.4	+2.5	384	0	0	31.2
expfrk12	+5.9	+7.0	+5.4	+2.5	384	0	0	31.2
expfrk11	+3.9	+4.5	+3.1	+1.0	256	0	0	23.9
eferk	+8.0	+9.3	+7.4	+4.1	256	1	0	123.9
liniger1	+2.8	+3.2	+1.9	+.5	128	1	1	80.6
liniger2	+7.7	+8.7	+7.0	+4.0	128	1	1	156.0
liniger3	+7.7	+8.7	+7.0	+4.0	128	1	1	156.8

$$h1 = +.78125 \cdot 10^{-1} \quad h2 = +.15625 \cdot 10^{-0}$$

method	sd1	sd6	sd27	sd51	fev	jev	lu	time
bw.euler	+2.0	+2.3	+1.0	-.4	193	2	2	166.2
efsirk	+7.9	+9.0	+7.3	+4.3	384	1	2	207.0
expfrk22	+8.3	+9.6	+7.7	+4.3	768	0	0	56.8
expfrk21	+6.2	+7.3	+5.7	+2.7	576	0	0	47.5
expfrk12	+6.2	+7.3	+5.7	+2.7	576	0	0	46.0
expfrk11	+4.1	+4.7	+3.3	+1.2	384	0	0	36.4
eferk	+8.3	+9.6	+7.7	+4.3	384	1	0	183.9
liniger1	+3.0	+3.4	+2.1	+.8	192	1	2	127.3
liniger2	+7.9	+9.0	+7.3	+4.3	192	1	2	245.7
liniger3	+7.9	+9.0	+7.3	+4.3	192	1	2	245.5

Opmerkingen:

De analytische oplossing luidt

$${}_k y(x) = \frac{1}{52} \sum_{i=1}^{51} 2 \sin \frac{ik\pi}{52} \sin \frac{i\pi}{52} \exp(-2x+2x \cos \frac{i\pi}{52}), \quad k=1, \dots, 51.$$

Met behulp van deze uitdrukking voor ${}_k y(x)$ is in dubbele lengte de referentie-oplossing in $x = 10$ en $x = 20$ berekend. Aangezien alle eigenwaarden op het interval $(-4,0)$ liggen, kunnen we niet direkt spreken van een stijf stelsel differentiaalvergelijking. Met een betrekkelijk grote integratiestap kan dan ook een redelijke precisie worden bereikt.

Verder zien we dat bij dezelfde $h2$ halvering van $h1$ bijna geen invloed heeft.

Voorbeeld 6

Differentiaalvergelijking: (zie [4])

$$y' = \begin{bmatrix} -1800 & 900 & 0 & \dots & 0 & \dots & 0 \\ 1 & -2 & 1 & 0 & \dots & & 0 \\ 0 & 1 & -2 & 1 & & & \\ \vdots & & & \ddots & \ddots & \ddots & \\ 0 & & & & 1 & -2 & 1 \\ 0 & & & & 0 & 1000 & -2000 \end{bmatrix}_{9 \times 9} y + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1000 \end{bmatrix}.$$

Beginvoorwaarden:

$$y(0) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Integratie-interval: $[0, 20]$.

Referentie-oplossing:

$$\begin{aligned} {}_1y(1/32) &= .1093223263520_{10^{-14}}, & {}_1y(20) &= .7760940921150_{10^{-1}}, \\ {}_4y(1/32) &= .1093753461966_{10^{-9}}, & {}_4y(20) &= .3258529791224_{10^{-0}}, \\ {}_7y(1/32) &= .2281173208623_{10^{-3}}, & {}_7y(20) &= .6372762332849_{10^{-0}}, \\ {}_9y(1/32) &= .5073927362337_{10^{-0}}, & {}_9y(20) &= .8775679605260_{10^{-0}}. \end{aligned}$$

Stuurparameters: $\delta = 2000.500375314$, $\delta^2 = 1800.500417052$, $\phi = \pi$.

results for integration until $x = +.31250_{10^{-1}}$

$h1 = +.97656_{10^{-3}}$

method	sd1	sd4	sd7	sd9	fev	jev	lu	time
bw.euler	+0	+5	+1.5	+5.0	33	1	1	3.7
efsirk	+4.1	+5.0	+8.8	+12.0	64	1	1	3.6
expfrk23	+2.9	+3.5	+6.0	+9.7	160	0	0	2.2
expfrk21	+1.9	+2.5	+4.9	+8.6	96	0	0	1.5
expfrk12	+1.0	+1.3	+2.3	+5.8	96	0	0	1.5
expfrk11	+6	+9	+1.9	+5.3	64	0	0	1.1
eferk	+3.5	+4.4	+8.1	+11.7	64	1	0	3.5
liniger1	+6	+1.0	+2.0	+5.5	32	1	1	2.7
liniger2	+2.9	+3.5	+6.0	+9.7	32	1	1	4.1
liniger3	+4.1	+5.0	+8.8	+11.7	32	1	1	4.1

$h1 = +.48828_{10^{-3}}$

method	sd1	sd4	sd7	sd9	fev	jev	lu	time
bw.euler	+4	+8	+1.8	+5.3	65	1	1	7.1
efsirk	+5.3	+6.2	+10.0	+12.0	128	1	1	6.6
expfrk23	+8.2	+10.3	+12.0	+11.7	320	0	0	4.2
expfrk21	+3.9	+4.7	+8.5	+12.0	192	0	0	2.9
expfrk12	+3.9	+4.7	+8.5	+12.0	192	0	0	2.9
expfrk11	+2.1	+2.6	+5.1	+8.7	128	0	0	2.2
eferk	+5.9	+7.3	+11.4	+12.0	128	1	0	6.6
liniger1	+1.2	+1.6	+2.6	+6.1	64	1	1	5.0
liniger2	+5.3	+6.2	+10.0	+12.0	64	1	1	9.3
liniger3	+5.3	+6.2	+10.0	+12.0	64	1	1	7.9

$h1 = +.24414_{10^{-3}}$

method	sd1	sd4	sd7	sd9	fev	jev	lu	time
bw.euler	+7	+1.1	+2.1	+5.6	129	1	1	14.4
efsirk	+6.5	+7.4	+12.0	+11.7	256	1	1	13.7
expfrk23	+9.7	+10.9	+12.0	+12.0	640	0	0	8.6
expfrk21	+4.8	+5.6	+9.4	+12.0	384	0	0	5.7
expfrk12	+4.8	+5.6	+9.4	+12.0	384	0	0	5.7
expfrk11	+2.7	+3.2	+5.7	+9.3	256	0	0	5.1
eferk	+7.1	+8.5	+11.3	+11.7	256	1	0	13.2
liniger1	+1.8	+2.2	+3.2	+6.7	128	1	1	9.9
liniger2	+6.5	+7.4	+11.2	+11.7	128	1	1	15.7
liniger3	+6.5	+7.4	+11.2	+11.7	128	1	1	14.9

results for integration until $x = +.20000_{10}+2$

$h1 = +.97656_{10}-3$ $h2 = +.15601_{10}-0$

method	sd1	sd4	sd7	sd9	fev	jev	lu	time
bw.euler	+2.3	+2.4	+2.8	+3.4	161	2	2	18.5
efsirk	+7.8	+7.8	+7.9	+8.5	320	1	2	16.2
expfrk23	-12.0	-12.0	-12.0	-12.0	800	0	0	11.4
expfrk21	-12.0	-12.0	-12.0	-12.0	480	0	0	7.1
expfrk12	-12.0	-12.0	-12.0	-12.0	480	0	0	7.1
expfrk11	-12.0	-12.0	-12.0	-12.0	320	0	0	5.4
eferk	-12.0	-12.0	-12.0	-12.0	320	1	0	17.9
liniger1	+2.3	+2.4	+2.8	+3.4	160	1	2	12.3
liniger2	+4.7	+4.7	+5.1	+5.7	160	1	2	19.1
liniger3	+7.8	+7.8	+7.9	+8.5	160	1	2	19.8

$h1 = +.48828_{10}-3$ $h2 = +.15601_{10}-0$

method	sd1	sd4	sd7	sd9	fev	jev	lu	time
bw.euler	+2.3	+2.4	+2.8	+3.4	193	2	2	22.1
efsirk	+7.8	+7.8	+7.9	+8.5	384	1	2	20.0
expfrk23	-12.0	-12.0	-12.0	-12.0	960	0	0	13.7
expfrk21	-12.0	-12.0	-12.0	-12.0	576	0	0	8.7
expfrk12	-12.0	-12.0	-12.0	-12.0	576	0	0	8.8
expfrk11	-12.0	-12.0	-12.0	-12.0	384	0	0	7.4
eferk	-12.0	-12.0	-12.0	-12.0	384	1	0	20.1
liniger1	+2.3	+2.4	+2.8	+3.4	192	1	2	15.4
liniger2	+4.7	+4.7	+5.1	+5.7	192	1	2	23.3
liniger3	+7.8	+7.8	+7.9	+8.5	192	1	2	22.3

$h1 = +.48828_{10}-3$ $h2 = +.78003_{10}-1$

method	sd1	sd4	sd7	sd9	fev	jev	lu	time
bw.euler	+2.6	+2.7	+3.1	+3.7	321	2	2	35.2
efsirk	+8.4	+8.7	+10.5	+11.7	640	1	2	33.2
expfrk23	-12.0	-12.0	-12.0	-12.0	1600	0	0	20.8
expfrk21	-12.0	-12.0	-12.0	-12.0	960	0	0	15.0
expfrk12	-12.0	-9.8	-.1	+.9	960	0	0	14.2
expfrk11	-12.0	-12.0	-12.0	-12.0	640	0	0	11.6
eferk	-12.0	-12.0	-12.0	-12.0	640	1	0	34.6
liniger1	+2.6	+2.7	+3.1	+3.7	320	1	2	23.8
liniger2	+5.3	+5.3	+5.7	+6.3	320	1	2	38.4
liniger3	+8.5	+8.7	+10.5	+10.4	320	1	2	37.5

$$h1 = +.24414_{10} - 3 \quad h2 = +.78003_{10} - 1$$

method	sd1	sd4	sd7	sd9	fev	jev	lu	time
bw.euler	+2.6	+2.7	+3.1	+3.7	385	2	2	42.5
efsirk	+8.5	+8.7	+10.6	+11.7	768	1	2	39.5
expfrk23	-12.0	-12.0	-12.0	-12.0	1920	0	0	25.0
expfrk21	-12.0	-12.0	-12.0	-12.0	1152	0	0	16.9
expfrk12	-12.0	-8.7	+1.3	+3.2	1152	0	0	16.9
expfrk11	-12.0	-12.0	-12.0	-12.0	768	0	0	12.9
eferk	-12.0	-12.0	-12.0	-12.0	768	1	0	40.2
liniger1	+2.6	+2.7	+3.1	+3.7	384	1	2	28.4
liniger2	+5.3	+5.3	+5.7	+6.3	384	1	2	43.8
liniger3	+8.5	+8.7	+10.5	+10.4	384	1	2	45.1

Opmerkingen:

We hebben de referentie-oplossing op twee manieren bepaald: namelijk door de eigenwaarden en eigenvectoren van het stelsel numeriek te berekenen, en door numerieke integratie met Liniger 2, waarbij een zeer kleine stap gebruikt werd (als kleinste waarde van $h1$ $1/16384$, en van $h2$ $1/128$). Extrapolatie van de op deze manier verkregen resultaten leverde de referentie-oplossing op in $xe1$; in dit punt was de oplossing met behulp van eigenvectoren door cijferverlies niet nauwkeurig. Voor de referentie-oplossing in $xe2$ hebben we de met eigenvectoren berekende waarde genomen.

Uit de tabellen voor $xe2$ blijkt dat de methoden $\text{expfrk } \langle i \rangle \langle j \rangle$ en eferk instabiel zijn bij deze staplengten. Dit wordt veroorzaakt doordat $-\delta_2$ niet in het stabiliteitsgebied van deze methoden ligt, omdat steeds in het punt $-\delta_2$ aangepast werd. We vermijden deze instabiliteit echter door alternierend in $-\delta_1$ en $-\delta_2$ aan te passen, zoals uit de volgende resultaten blijkt:

results for integration until $x = +.20000_{10}+2$

$h1 = +.97656_{10}-3$ $h2 = +.15601_{10}-0$

method	sd1	sd4	sd7	sd9	fev	jev	lu	time
expfrk23	+4.7	+4.7	+5.1	+5.4	800	0	0	13.2
expfrk21	+4.6	+4.7	+5.1	+4.6	480	0	0	8.2
expfrk12	+2.3	+2.4	+2.8	+3.4	480	0	0	9.3
expfrk11	+2.3	+2.4	+2.8	+3.4	320	0	0	6.4
eferk	-10.8	-10.5	-12.0	-12.0	320	1	0	17.8

$h1 = +.48828_{10}-3$ $h2 = +.15601_{10}-0$

method	sd1	sd4	sd7	sd9	fev	jev	lu	time
expfrk23	+4.6	+4.7	+5.1	+6.0	960	0	0	15.2
expfrk21	+4.6	+4.7	+5.1	+4.3	576	0	0	9.5
expfrk12	+2.3	+2.4	+2.8	+3.4	576	0	0	10.7
expfrk11	+2.3	+2.4	+2.8	+3.4	384	0	0	7.6
eferk	-5.2	-6.0	-12.0	-12.0	384	1	0	21.3

$h1 = +.48828_{10}-3$ $h2 = +.78003_{10}-1$

method	sd1	sd4	sd7	sd9	fev	jev	lu	time
expfrk23	+5.3	+5.4	+5.7	+6.3	1600	0	0	27.2
expfrk21	+5.3	+5.3	+5.7	+5.0	960	0	0	16.1
expfrk12	+2.7	+2.8	+3.1	+3.7	960	0	0	17.4
expfrk11	+2.6	+2.7	+3.1	+3.7	640	0	0	12.4
eferk	+8.7	+8.4	+8.5	+5.1	640	1	0	35.7

$h1 = +.24414_{10}-3$ $h2 = +.78003_{10}-1$

method	sd1	sd4	sd7	sd9	fev	jev	lu	time
expfrk23	+5.3	+5.4	+5.7	+6.3	1920	0	0	31.4
expfrk21	+5.3	+5.3	+5.7	+5.6	1152	0	0	19.7
expfrk12	+2.7	+2.8	+3.1	+3.7	1152	0	0	20.3
expfrk11	+2.6	+2.7	+3.1	+3.7	768	0	0	14.6
eferk	+8.7	+8.4	+8.5	+5.2	768	1	0	41.3

Voorbeeld 7

Differentiaalvergelijking: (zie [5])

$$y_1' = -1000y_1(y_1 + y_2 - 1.999987)$$

$$y_2' = -2500y_2(y_1 + y_2 - 2).$$

Beginvoorwaarden:

$$y(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Integratie-interval: [0,50].

Referentie-oplossing:

$$y_1\left(\frac{1}{64}\right) = .999853854436, \quad y_1(50) = .5976546988,$$

$$y_2\left(\frac{1}{64}\right) = 1.00014243203, \quad y_2(50) = 1.4023434075.$$

Stuurparameters: $\phi = \pi$, δ_1 en δ_2 werden iedere stap als volgt berekend:

$$\delta_1 = \begin{cases} \max(|\delta_1|, |\delta_2|), & \text{als } \operatorname{Re} \delta_1 < 0 \text{ of } \operatorname{Re} \delta_2 < 0, \\ 0 & , \text{als } \operatorname{Re} \delta_1 \geq 0 \text{ en } \operatorname{Re} \delta_2 \geq 0, \end{cases}$$

$$\delta_2 = \begin{cases} \min(|\delta_1|, |\delta_2|), & \text{als } \operatorname{Re} \delta_1 < 0 \text{ en } \operatorname{Re} \delta_2 < 0, \\ 0 & , \text{als } \operatorname{Re} \delta_1 \geq 0 \text{ of } \operatorname{Re} \delta_2 \geq 0, \end{cases}$$

waarin δ_1 en δ_2 de eigenwaarden van de jacobiaan zijn.

results for integration until $x = +.15625_{10^{-1}}$

$h1 = +.48828_{10^{-3}}$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+9.9	+9.9	33	1	1	.9
efsirk	+10.9	+11.0	64	32	32	1.7
expfrk23	+11.0	+11.3	160	0	0	1.3
expfrk22	+11.1	+11.0	128	0	0	1.0
expfrk13	+10.9	+11.0	128	0	0	1.1
expfrk11	+10.2	+10.2	64	0	0	.6
eferk	+11.0	+11.0	64	32	0	1.0
liniger1	+10.5	+10.4	32	32	32	1.5
liniger2	+11.2	+11.3	32	32	32	2.0
liniger3	+11.2	+11.3	32	32	32	2.0

$h1 = +.24414_{10^{-3}}$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+10.2	+10.2	65	1	1	1.8
efsirk	+10.8	+10.9	128	64	64	3.6
expfrk23	+11.0	+11.1	320	0	0	1.6
expfrk22	+11.1	+11.3	256	0	0	1.3
expfrk13	+11.1	+11.3	256	0	0	1.3
expfrk11	+11.1	+11.3	128	0	0	.9
eferk	+11.0	+11.0	128	64	0	1.8
liniger1	+11.4	+11.3	64	64	64	2.9
liniger2	+10.8	+10.8	64	64	64	4.7
liniger3	+10.8	+10.8	64	64	64	4.0

$h1 = +.12207_{10^{-3}}$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+10.6	+10.7	129	1	1	3.4
efsirk	+11.1	+11.3	256	128	128	7.2
expfrk23	+11.3	+11.4	640	0	0	3.2
expfrk22	+11.1	+11.3	512	0	0	2.7
expfrk13	+11.1	+11.3	512	0	0	2.6
expfrk11	+11.0	+11.1	256	0	0	1.8
eferk	+11.3	+11.7	256	128	0	4.4
liniger1	+11.0	+11.0	128	128	128	6.0
liniger2	+12.0	+11.7	128	128	128	8.2
liniger3	+12.0	+11.7	128	128	128	8.0

results for integration until $x = +.50000_{10}+2$

$h1 = +.48828_{10}-3$ $h2 = +.39050_{10}-0$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+3.1	+3.5	231	2	2	5.6
efsirk	-12.0	-12.0	320	160	160	9.6
expfrk23	+5.8	+5.3	800	0	0	6.1
expfrk22	-12.0	-12.0	640	0	0	3.7
expfrk13	+3.1	+3.5	640	0	0	5.2
expfrk11	+3.1	+3.5	320	0	0	2.9
eferk	-12.0	-12.0	320	160	0	5.4
liniger1	+3.1	+3.5	288	160	160	9.5
liniger2	+5.9	+6.2	288	160	160	12.5
liniger3	+7.5	+7.8	288	288	288	17.5

$h1 = +.24414_{10}-3$ $h2 = +.39050_{10}-0$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+3.1	+3.5	263	2	2	6.3
efsirk	-12.0	-12.0	384	192	192	10.6
expfrk23	+5.8	+5.5	960	0	0	7.2
expfrk22	-12.0	-12.0	768	0	0	4.2
expfrk13	+3.1	+3.5	768	0	0	5.5
expfrk11	+3.1	+3.5	384	0	0	3.2
eferk	-12.0	-12.0	384	192	0	5.4
liniger1	+3.1	+3.5	320	192	192	11.7
liniger2	+5.9	+6.2	320	192	192	14.4
liniger3	+7.5	+7.8	320	320	320	19.4

$h1 = +.24414_{10}-3$ $h2 = +.19525_{10}-0$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+3.4	+3.8	321	2	2	8.6
efsirk	+8.3	+8.7	640	320	320	18.0
expfrk23	+7.0	+6.5	1600	0	0	11.4
expfrk22	+6.7	+6.5	1280	0	0	9.0
expfrk13	+3.4	+3.8	1280	0	0	10.5
expfrk11	+3.4	+3.8	640	0	0	5.4
eferk	+8.3	+8.7	640	320	0	9.8
liniger1	+3.4	+3.8	576	320	320	19.9
liniger2	+6.5	+6.9	576	320	320	24.8
liniger3	+7.5	+7.8	576	576	576	35.0

$$h1 = +.12207_{10} - 3 \quad h2 = +.19525_{10} - 0$$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+3.4	+3.8	385	2	2	10.2
efsirk	+8.2	+8.7	768	384	384	21.6
expfrk23	+6.5	+6.2	1920	0	0	12.7
expfrk22	+6.7	+6.5	1536	0	0	10.3
expfrk13	+3.4	+3.8	1536	0	0	11.6
expfrk11	+3.4	+3.8	768	0	0	6.3
eferk	+8.2	+8.5	768	384	0	11.5
liniger1	+3.4	+3.8	640	384	384	22.6
liniger2	+6.5	+6.9	640	384	384	29.5
liniger3	+7.5	+7.8	640	640	640	40.3

Opmerkingen:

De referentie-oplossing werd bepaald met behulp van MULTISTEP. Uit de 4^e en 5^e tabel blijkt dat efsirk en eferk gevoelig zijn voor een sterk veranderende jacobiaan. Integreren we bij deze procedures namelijk met een te grote h2 dan zal dit aanleiding geven tot instabiliteit omdat de differentiaalvergelijking dan, lokaal gezien, niet meer als lineair te beschouwen is.

Voorbeeld 8

Differentiaalvergelijking: (zie [3])

$$y_1' = .2(y_2 - y_1)$$

$$y_2' = 10y_1 - (60 - .125y_3)y_2 + .125y_3$$

$$y_3' = 1.$$

Beginvoorwaarden:

$$y(0) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Integratie-interval: [0,400].

Referentie-oplossing:

$$y_1(1) = .1910911314409_{10}^{-3}, \quad y_1(400) = 22.24222011,$$

$$y_2(1) = .2083621007881_{10}^{-2}, \quad y_2(400) = 27.11071335,$$

$$y_3(1) = 1, \quad y_3(400) = 400.$$

Stuurparameters: $\phi = \pi$, δ en δ^2 werden op dezelfde manier berekend als in voorbeeld 7.

results for integration until $x = +.10000_{10}+1$

$h1 = +.31250_{10}- 1$

method	sd1	sd2	sd3	fev	jev	lu	time
bw.euler	+1.5	+3.4	+12.0	35	1	1	1.3
efsirk	+6.2	+6.3	+12.0	64	32	32	4.3
expfrk11	+1.9	+3.7	+12.0	64	0	0	.7
expfrk12	+2.4	+4.2	+12.0	96	0	0	1.0
expfrk13	+2.9	+4.7	+12.0	128	0	0	1.2
expfrk21	+4.8	+6.8	+12.0	96	0	0	.9
eferk	+6.2	+6.3	+12.0	64	32	0	1.4
liniger1	+2.1	+3.9	+12.0	64	32	32	3.1
liniger2	+5.8	+6.4	+12.0	64	32	32	4.2
liniger3	+6.2	+8.3	+12.0	64	64	64	6.2

$h1 = +.15625_{10}- 1$

method	sd1	sd2	sd3	fev	jev	lu	time
bw.euler	+1.8	+3.7	+12.0	67	1	1	2.5
efsirk	+6.2	+7.3	+12.0	128	64	64	6.0
expfrk11	+4.9	+6.4	+12.0	128	0	0	1.8
expfrk12	+6.3	+7.0	+12.0	192	0	0	1.3
expfrk13	+6.3	+7.3	+12.0	256	0	0	1.6
expfrk21	+6.3	+7.0	+12.0	192	0	0	1.4
eferk	+6.2	+7.2	+12.0	128	64	0	2.6
liniger1	+2.7	+4.5	+12.0	128	64	64	6.3
liniger2	+6.1	+7.0	+12.0	128	64	64	8.5
liniger3	+6.2	+8.3	+12.0	128	128	128	13.0

$h1 = +.78125_{10}- 2$

method	sd1	sd2	sd3	fev	jev	lu	time
bw.euler	+2.1	+4.0	+12.0	129	1	1	4.7
efsirk	+6.2	+9.1	+12.0	256	128	128	12.0
expfrk11	+5.4	+7.1	+12.0	256	0	0	2.1
expfrk12	+6.2	+7.6	+12.0	384	0	0	2.6
expfrk13	+6.2	+7.8	+12.0	512	0	0	3.2
expfrk21	+6.2	+7.6	+12.0	384	0	0	3.4
eferk	+6.2	+8.9	+12.0	256	128	0	4.9
liniger1	+3.2	+5.1	+11.7	256	128	128	12.3
liniger2	+6.2	+7.7	+12.0	256	128	128	17.7
liniger3	+6.2	+8.3	+12.0	256	256	256	24.5

results for integration until $x = +.40000_{10}+3$

$h1 = +.31250_{10}-1$ $h2 = +.77930_{10}-0$

method	sd1	sd2	sd3	fev	jev	lu	time
bw.euler	+2.1	+2.1	+12.0	1338	12	12	39.9
efsirk	+4.8	+3.2	+12.0	1088	544	544	51.2
expfrk11	+2.0	+2.1	+12.0	1088	0	0	11.0
expfrk12	+2.1	+2.2	+12.0	1632	0	0	15.4
expfrk13	+2.2	+2.3	+12.0	2176	0	0	19.8
expfrk21	+4.4	+3.6	+12.0	1632	0	0	13.9
eferk	+4.8	+3.2	+12.0	1088	544	0	22.7
liniger1	+2.0	+2.1	+12.0	1088	544	544	52.0
liniger2	+3.6	+3.6	+12.0	1088	544	544	70.8
liniger3	+7.0	+6.3	+12.0	1088	1088	1088	103.9

$h1 = +.15625_{10}-1$ $h2 = +.77930_{10}-0$

method	sd1	sd2	sd3	fev	jev	lu	time
bw.euler	+2.1	+2.1	+12.0	1370	12	12	41.0
efsirk	+4.8	+3.2	+12.0	1152	576	576	53.4
expfrk11	+2.0	+2.1	+12.0	1152	0	0	11.4
expfrk12	+2.1	+2.2	+12.0	1728	0	0	15.8
expfrk13	+2.2	+2.3	+12.0	2304	0	0	20.2
expfrk21	+4.4	+3.6	+12.0	1728	0	0	14.4
eferk	+4.8	+3.2	+12.0	1152	576	0	23.8
liniger1	+2.0	+2.1	+12.0	1152	576	576	55.0
liniger2	+3.6	+3.6	+12.0	1152	576	576	74.9
liniger3	+7.0	+6.3	+12.0	1152	1152	1152	110.0

$h1 = +.15625_{10}-1$ $h2 = +.38965_{10}-0$

method	sd1	sd2	sd3	fev	jev	lu	time
bw.euler	+2.2	+2.3	+12.0	2145	6	6	67.0
efsirk	+5.9	+4.2	+12.0	2176	1088	1088	100.7
expfrk11	+2.4	+2.5	+12.0	2176	0	0	21.7
expfrk12	+2.6	+2.7	+12.0	3264	0	0	30.2
expfrk13	+2.9	+2.9	+12.0	4352	0	0	38.7
expfrk21	+5.0	+5.2	+12.0	3264	0	0	27.4
eferk	+5.9	+4.2	+12.0	2176	1088	0	45.2
liniger1	+2.5	+2.5	+12.0	2176	1088	1088	103.9
liniger2	+4.3	+4.2	+12.0	2176	1088	1088	140.9
liniger3	+8.5	+7.3	+12.0	2176	2176	2176	207.8

$h1 = +.78125_{10} - 2$ $h2 = +.38965_{10} - 0$

method	sd1	sd2	sd3	fev	jev	lu	time
bw.euler	+2.2	+2.3	+12.0	2207	6	6	68.2
efsirk	+5.9	+4.2	+12.0	2304	1152	1152	106.7
expfrk11	+2.4	+2.5	+12.0	2304	0	0	22.8
expfrk12	+2.6	+2.7	+12.0	3456	0	0	31.5
expfrk13	+2.9	+2.9	+12.0	4608	0	0	40.3
expfrk21	+5.0	+5.2	+12.0	3456	0	0	28.7
eferk	+5.9	+4.2	+12.0	2304	1152	0	47.6
liniger1	+2.5	+2.5	+12.0	2304	1152	1152	109.9
liniger2	+4.3	+4.2	+12.0	2304	1152	1152	148.4
liniger3	+8.5	+7.3	+12.0	2304	2304	2304	219.9

Opmerkingen:

Ook van dit probleem werd de referentie-oplossing bepaald met behulp van MULTISTEP.

Het verloop van de resultaten bij dit probleem verschilt in wezen niet van dat bij vorige voorbeelden.

Voorbeeld 9

Differentiaalvergelijking: (zie [6])

$$y_1' = .01 - [1+(y_1+1000)(y_1+1)](.01+y_1+y_2)$$

$$y_2' = .01 - (1+y_2^2)(.01+y_1+y_2).$$

Beginvoorwaarden:

$$y(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Integratie-interval: [0,100].

Referentie-oplossing:

$$y_1\left(\frac{1}{32}\right) = -.10281693097_{10}^{-1}, \quad y_1(100) = -.9916420658355,$$

$$y_2\left(\frac{1}{32}\right) = .30186013_{10}^{-3}, \quad y_2(100) = .9833363542311.$$

Stuurparameters: $\phi = \pi$, δ en δ^2 werden op dezelfde manier berekend als in voorbeeld 7.

results for integration until $x = +.31250_{10^{-1}}$

$h1 = +.97656_{10^{-3}}$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+5.8	+4.3	44	1	1	1.1
efsirk	+6.9	+5.3	64	32	32	1.8
expfrk23	+6.5	+4.9	160	0	0	.9
expfrk21	+8.2	+6.6	96	0	0	.6
expfrk22	+6.4	+4.8	128	0	0	.8
expfrk12	+8.2	+6.6	96	0	0	.6
eferk	+7.1	+5.5	64	32	0	.9
liniger1	+6.2	+4.6	66	32	32	2.1
liniger2	+6.4	+4.8	66	32	32	2.7
liniger3	+7.4	+5.8	66	66	66	3.9

$h1 = +.48828_{10^{-3}}$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+6.0	+4.5	79	1	1	2.0
efsirk	+7.7	+6.2	128	64	64	3.7
expfrk23	+7.0	+5.5	320	0	0	1.7
expfrk21	+7.2	+5.7	192	0	0	1.2
expfrk22	+7.0	+5.5	256	0	0	1.5
expfrk12	+7.2	+5.7	192	0	0	1.2
eferk	+7.8	+6.3	128	64	0	1.9
liniger1	+6.7	+5.2	130	64	64	4.1
liniger2	+6.8	+5.3	130	64	64	5.4
liniger3	+8.8	+7.0	130	130	130	7.6

$h1 = +.24414_{10^{-3}}$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+6.3	+4.8	148	1	1	3.7
efsirk	+8.5	+7.1	256	128	128	7.3
expfrk23	+7.6	+6.1	640	0	0	3.4
expfrk21	+7.7	+6.2	384	0	0	2.4
expfrk22	+7.6	+6.1	512	0	0	2.9
expfrk12	+7.7	+6.2	384	0	0	2.4
eferk	+8.5	+7.2	256	128	0	3.7
liniger1	+7.3	+5.8	257	128	128	8.2
liniger2	+7.4	+5.8	257	128	128	10.6
liniger3	+8.9	+8.2	257	257	257	15.0

results for integration until $x = +.10000_{10}+3$

$h1 = +.97656_{10}-3$ $h2 = +.39050_{10}-0$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+3.7	+3.5	325	4	4	8.3
efsirk	+4.5	+5.2	576	288	288	16.1
expfrk23	+5.0	+5.1	1440	0	0	11.2
expfrk21	+4.9	+5.1	864	0	0	6.7
expfrk22	+5.8	+5.1	1152	0	0	9.0
expfrk12	+3.5	+3.4	864	0	0	7.4
eferk	+4.5	+5.2	576	288	0	9.2
liniger1	+3.4	+3.4	578	288	288	18.4
liniger2	+4.6	+4.6	578	288	288	25.3
liniger3	+5.6	+6.0	578	578	578	34.5

$h1 = +.48828_{10}-3$ $h2 = +.39050_{10}-0$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+3.7	+3.5	360	4	4	9.2
efsirk	+4.5	+5.2	640	320	320	17.9
expfrk23	+5.0	+5.1	1600	0	0	12.1
expfrk21	+4.9	+5.1	960	0	0	7.3
expfrk22	+5.8	+5.1	1280	0	0	9.7
expfrk12	+3.5	+3.4	960	0	0	8.0
eferk	+4.5	+5.2	640	320	0	10.1
liniger1	+3.4	+3.4	642	320	320	20.4
liniger2	+4.6	+4.6	642	320	320	26.5
liniger3	+5.6	+6.0	642	642	642	37.4

$h1 = +.48828_{10}-3$ $h2 = +.19525_{10}-0$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+3.1	+3.3	609	2	2	15.7
efsirk	+5.4	+6.2	1152	576	576	33.0
expfrk23	+5.4	+5.6	2880	0	0	22.7
expfrk21	+5.9	+5.7	1728	0	0	13.4
expfrk22	+5.5	+5.7	2304	0	0	17.9
expfrk12	+4.0	+3.9	1728	0	0	14.8
eferk	+5.4	+6.2	1152	576	0	18.4
liniger1	+3.8	+3.8	1154	576	576	36.7
liniger2	+5.1	+5.2	1154	576	576	49.9
liniger3	+6.8	+7.0	1154	1154	1154	69.1

$$h1 = +.24414_{10^{-3}} \quad h2 = +.19525_{10^{-0}}$$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+3.1	+3.3	678	2	2	17.5
efsirk	+5.4	+6.2	1280	640	640	36.0
expfrk23	+5.4	+5.6	3200	0	0	25.0
expfrk21	+5.9	+5.7	1920	0	0	14.9
expfrk22	+5.5	+5.7	2560	0	0	19.3
expfrk12	+4.0	+3.9	1920	0	0	16.0
eferk	+5.4	+6.2	1280	640	0	20.3
liniger1	+3.8	+3.8	1281	640	640	40.7
liniger2	+5.1	+5.2	1281	640	640	54.0
liniger3	+6.8	+7.0	1281	1281	1281	75.4

Opmerkingen:

De referentie-oplossing van dit probleem werd met MULTISTEP bepaald. Opvallend is dat op het begin-interval expfrk 12 minstens zulke nauwkeurige resultaten geeft als de methoden expfrk 2<j>, hetgeen tot nu toe niet eerder is voorgekomen.

Voorbeeld 10

Differentiaalvergelijking: (zie [7])

$$y_1' = -K_1 y_1$$

$$y_2' = K_1 y_1 + K_{11} K_{14} y_4 + K_{19} K_{14} y_5 - K_3 y_2 y_3 - K_{15} y_2 y_{12} - K_2 y_2$$

$$y_3' = K_2 y_2 - K_5 y_3 - K_3 y_2 y_3 - K_7 y_{10} y_3 + K_{11} K_{14} y_4 + K_{12} K_{14} y_6$$

$$y_4' = K_3 y_2 y_3 - K_{11} K_{14} y_4 - K_4 y_4$$

$$y_5' = K_{15} y_2 y_{12} - K_{19} K_{14} y_5 - K_{16} y_5$$

$$y_6' = K_7 y_{10} y_3 - K_{12} K_{14} y_6 - K_8 y_6$$

$$y_7' = K_{17} y_{10} y_{12} - K_{20} K_{14} y_7 - K_{18} y_7$$

$$y_8' = K_9 y_{10} - K_{13} K_{14} y_8 - K_{10} y_8$$

$$y_9' = K_4 y_4 + K_{16} y_5 + K_8 y_6 + K_{18} y_7$$

$$y_{10}' = K_5 y_3 + K_{12} K_{14} y_6 + K_{20} K_{14} y_7 + K_{13} K_{14} y_8 - K_7 y_{10} y_3$$

$$- K_{17} y_{10} y_{12} - K_6 y_{10} - K_9 y_{10}$$

$$y_{11}' = K_{10} y_8$$

$$y_{12}' = K_6 y_{10} + K_{19} K_{14} y_5 + K_{20} K_{14} y_7 - K_{15} y_2 y_{12} - K_{17} y_{10} y_{12},$$

met:

$$K_2 = K_6 = 10.0, K_1 = K_5 = 0.1, K_{10} = 5.0,$$

$$K_{14} = 30.0, K_4 = K_{16} = K_8 = K_{18} = 2.5,$$

$$K_3 = K_7 = K_9 = K_{11} = K_{19} = K_{12} = K_{20} = K_{13} = 50.0,$$

$$K_{15} = K_{17} = 100.0.$$

(Merk op: $\sum_{i=1}^{12} y_i + y_4 + y_5 + y_6 + y_7 + y_9 = 1$)

Beginvoorwaarden:

$$y(0) = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Integratie-interval: $[0, 50]$.

Referentie-oplossing:

$$y_3\left(\frac{1}{64}\right) = .115825313_{10^{-3}}, \quad y_3(50) = .334500768_{10^{-1}},$$

$$y_5\left(\frac{1}{64}\right) = .179687_{10^{-12}}, \quad y_5(50) = .407994_{10^{-5}},$$

$$y_9\left(\frac{1}{64}\right) = .476247_{10^{-10}}, \quad y_9(50) = .149109210_{10^{-1}},$$

$$y_{12}\left(\frac{1}{64}\right) = .2269204_{10^{-8}}, \quad y_{12}(50) = .914169996.$$

Stuurparameters: $\delta = 1555.1617953$, $\delta^2 = 1502.5$, $\phi = \pi$.

results for integration until $x = +.15625_{10}-1$

$h1 = +.15625_{10}-1$

method	sd3	sd5	sd9	sd12	fev	jev	lu	time
bw.euler	+1	-1.3	-.8	-1.3	3	1	1	.9
efsirk	+3.1	-.5	+0	+7	2	1	1	2.5
expfrk12	+1	+0	+0	+0	3	0	0	.1
expfrk23	+1.4	+0	+0	+0	5	0	0	.2
expfrk22	+1.4	+0	+0	+0	4	0	0	.1
expfrk31	+2.7	+0	+7	+1	4	0	0	.1
eferk	+2.7	+0	+0	+1	2	1	0	.2
liniger1	+1	-1.2	-.8	-1.2	2	1	1	.8
liniger2	+1.5	-.7	-.5	-.6	2	1	1	1.2
liniger3	+3.3	+6	+6	+7	2	2	2	2.0

$h1 = +.78125_{10}-2$

method	sd3	sd5	sd9	sd12	fev	jev	lu	time
bw.euler	+4	-.8	-.4	-.8	5	1	1	1.2
efsirk	+4.0	+0	+7	+1.7	4	2	2	2.0
expfrk12	+5	+0	+1	+0	6	0	0	.2
expfrk23	+2.2	+5	+7	+5	10	0	0	.2
expfrk22	+2.1	+4	+6	+4	8	0	0	.2
expfrk31	+4.0	-.1	+4	+1.1	8	0	0	.2
eferk	+3.8	+1	+1.1	+1.1	4	2	0	.4
liniger1	+4	-.7	-.3	-.7	4	2	2	1.5
liniger2	+2.2	+2	+3	+2	4	2	2	2.3
liniger3	+4.3	+1.6	+1.6	+1.7	4	4	4	4.0

$h1 = +.39062_{10}-2$

method	sd3	sd5	sd9	sd12	fev	jev	lu	time
bw.euler	+7	-.4	-.0	-.3	9	1	1	1.7
efsirk	+5.0	+7	+1.6	+2.7	8	4	4	4.0
expfrk12	+9	+2	+4	+3	12	0	0	.3
expfrk23	+3.2	+1.2	+1.5	+1.3	20	0	0	.4
expfrk22	+2.9	+1.1	+1.4	+1.1	16	0	0	.3
expfrk31	+4.9	+9	+1.9	+2.1	16	0	0	.3
eferk	+4.7	+8	+1.7	+2.1	8	4	0	.7
liniger1	+8	-.2	+2	-.1	8	4	4	2.9
liniger2	+3.2	+1.6	+1.0	+1.2	8	4	4	5.6
liniger3	+5.3	+2.6	+2.6	+2.7	8	8	8	8.0

results for integration until $x = +.50000_{10}+2$

$h1 = +.15625_{10}-1$ $h2 = +.78101_{10}-0$

method	sd3	sd5	sd9	sd12	fev	jev	lu	time
bw.euler	-1.3	+0	+0	+0	195	66	66	65.5
efsirk	-12.0	-12.0	-12.0	-12.0	130	65	65	66.9
expfrk12	-12.0	-12.0	-12.0	-12.0	195	0	0	3.7
expfrk23	-12.0	-12.0	-12.0	-12.0	325	0	0	5.5
expfrk22	-12.0	-12.0	-12.0	-12.0	260	0	0	5.5
expfrk31	-12.0	-12.0	-12.0	-12.0	260	0	0	4.6
eferk	-12.0	-12.0	-12.0	-12.0	130	65	0	10.0
liniger1	+9	+7	+1.8	+2.2	130	65	65	49.0
liniger2	+2.7	+2.3	+3.4	+3.9	131	65	65	77.1
liniger3	+5.1	+4.3	+3.4	+4.9	130	130	130	135.9

$h1 = +.78125_{10}-2$ $h2 = +.78101_{10}-0$

method	sd3	sd5	sd9	sd12	fev	jev	lu	time
bw.euler	-1.3	+0	+0	+0	197	66	66	65.6
efsirk	-12.0	-12.0	-12.0	-12.0	132	66	66	67.9
expfrk12	-12.0	-12.0	-12.0	-12.0	198	0	0	3.8
expfrk23	-12.0	-12.0	-12.0	-12.0	330	0	0	5.6
expfrk22	-12.0	-12.0	-12.0	-12.0	264	0	0	4.7
expfrk31	-12.0	-12.0	-12.0	-12.0	264	0	0	4.6
eferk	-12.0	-12.0	-12.0	-12.0	132	66	0	11.1
liniger1	+9	+7	+1.8	+2.2	132	66	66	48.6
liniger2	+2.7	+2.3	+3.4	+3.9	133	66	66	79.1
liniger3	+5.1	+4.3	+3.4	+4.9	132	132	132	136.8

$h1 = +.78125_{10}-2$ $h2 = +.39050_{10}-0$

method	sd3	sd5	sd9	sd12	fev	jev	lu	time
bw.euler	+1.2	+1.0	+2.1	+2.5	300	2	2	39.2
efsirk	+5.4	+2.0	+4.2	+5.7	260	130	130	134.3
expfrk12	-12.0	-12.0	-12.0	-12.0	390	0	0	7.4
expfrk23	-12.0	-12.0	-12.0	-12.0	650	0	0	11.8
expfrk22	-12.0	-12.0	-12.0	-12.0	520	0	0	9.0
expfrk31	-12.0	-12.0	-12.0	-12.0	520	0	0	9.0
eferk	-12.0	-12.0	-12.0	-12.0	260	130	0	20.9
liniger1	+1.2	+1.0	+2.1	+2.5	260	130	130	96.5
liniger2	+3.3	+2.9	+4.1	+4.5	260	130	130	154.4
liniger3	+5.9	+5.2	+4.2	+5.8	260	260	260	271.1

$$h_1 = +.39062_{10^{-2}} \quad h_2 = +.39050_{10^{-0}}$$

method	sd3	sd5	sd9	sd12	fev	jev	lu	time
bw.euler	+1.2	+1.0	+2.1	+2.5	304	2	2	38.7
efsirk	+5.4	+2.0	+4.2	+5.7	264	132	132	136.3
expfrk12	-12.0	-12.0	-12.0	-12.0	396	0	0	8.4
expfrk23	-12.0	-12.0	-12.0	-12.0	660	0	0	11.0
expfrk22	-12.0	-12.0	-12.0	-12.0	528	0	0	9.2
expfrk31	-12.0	-12.0	-12.0	-12.0	528	0	0	10.1
eferk	-12.0	-12.0	-12.0	-12.0	264	132	0	20.1
liniger1	+1.2	+1.0	+2.1	+2.5	264	132	132	98.0
liniger2	+3.3	+2.9	+4.1	+4.5	264	132	132	156.7
liniger3	+5.9	+5.2	+4.2	+5.8	264	264	264	275.1

Opmerkingen:

De referentie-oplossing werd bepaald met behulp van EFSIRK en MULTI-STEP.

Voor delta en delta 2 namen we de maximale eigenwaarden van de jacobiaan in het punt $x = 0$.

De procedures expfrk<i><j> en eferk blijken voor alle gebruikte waarden van h_2 instabiel te zijn, terwijl efsirk slechts voor de kleinste waarde van h_2 bevredigende resultaten geeft.

We zien ook weer in dit voorbeeld dat, min of meer onafhankelijk van de oplossing in xe_1 , bij een geschikte h_2 gunstige resultaten kunnen worden verkregen in xe_2 .

Voorbeeld 11

Differentiaalvergelijking:

$$y_1' = -y_2$$

$$y_2' = y_1.$$

Beginvoorwaarden:

$$y(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Integratie-interval: $[0, 2\pi]$.

Referentie-oplossing:

$$y = \begin{pmatrix} \cos x \\ \sin x \end{pmatrix}.$$

Stuurparameters: $\phi = \pi$, $\delta_1 = \delta_2 = 0$.

results for integration until $x = +.62832_{10}+1$

$h1 = +.15708_{10}+ 1$

method	sd1	sd2	fev	jev	lu	time
bw.euler	-.0	+1.2	5	1	1	.2
efsirk	+3.0	+1.3	8	1	2	.2
expfrk11	-.6	-.5	8	0	0	.1
expfrk21	+.3	+.2	12	0	0	.1
expfrk31	+.6	+1.2	16	0	0	.1
expfrk41	+1.8	+1.1	20	0	0	1.4
eferk	+.6	+1.2	8	1	0	.2
liniger1	+.4	+.1	4	1	2	.2
liniger2	+3.0	+1.3	4	1	2	.2
liniger3	+3.0	+1.3	4	1	2	.2

$h1 = +.39270_{10}- 0$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+.2	+1.0	17	1	1	.6
efsirk	+7.7	+3.7	34	1	2	.5
expfrk11	+1.4	+.8	32	0	0	.2
expfrk21	+1.8	+2.3	48	0	0	.3
expfrk31	+3.4	+2.9	64	0	0	.3
expfrk41	+4.1	+4.6	80	0	0	.4
eferk	+3.4	+2.9	32	1	0	.4
liniger1	+2.5	+1.1	16	1	2	.4
liniger2	+7.7	+3.7	16	1	2	.6
liniger3	+7.7	+3.7	16	1	2	.6

$h1 = +.19635_{10}- 0$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+.3	+1.4	33	1	1	1.0
efsirk	+10.1	+4.9	66	1	2	.8
expfrk11	+2.3	+1.4	64	0	0	.4
expfrk21	+2.7	+3.5	96	0	0	.5
expfrk31	+4.9	+4.1	128	0	0	.7
expfrk41	+5.6	+6.4	160	0	0	.8
eferk	+4.9	+4.1	64	1	0	.9
liniger1	+3.7	+1.7	32	1	2	.8
liniger2	+10.1	+4.9	32	1	2	1.0
liniger3	+10.1	+4.9	32	1	2	1.0

$h1 = +.98175_{10} - 1$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+6	+1.8	65	1	1	2.0
efsirk	+11.4	+6.1	128	1	2	1.6
expfrk11	+3.2	+2.0	128	0	0	.9
expfrk21	+3.6	+4.7	192	0	0	1.1
expfrk31	+6.4	+5.3	256	0	0	1.3
expfrk41	+7.1	+8.2	320	0	0	1.6
eferk	+6.4	+5.3	128	1	0	1.7
liniger1	+4.9	+2.3	64	1	2	1.4
liniger2	+11.7	+6.1	64	1	2	2.0
liniger3	+11.7	+6.1	64	1	2	2.0

$h1 = +.49087_{10} - 1$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+8	+2.4	129	1	1	4.4
efsirk	+11.7	+7.3	258	1	2	2.8
expfrk11	+4.0	+2.6	256	0	0	1.5
expfrk21	+4.5	+5.9	384	0	0	1.9
expfrk31	+7.9	+6.5	512	0	0	2.4
expfrk41	+8.6	+10.2	640	0	0	2.8
eferk	+7.9	+6.5	256	1	0	3.1
liniger1	+6.1	+2.9	128	1	2	2.5
liniger2	+11.4	+7.3	128	1	2	3.5
liniger3	+11.4	+7.3	128	1	2	3.5

$h1 = +.24544_{10} - 1$

method	sd1	sd2	fev	jev	lu	time
bw.euler	+1.1	+2.9	257	1	1	9.1
efsirk	+12.0	+8.5	514	1	2	7.3
expfrk11	+4.9	+3.2	514	0	0	4.0
expfrk21	+5.4	+7.1	771	0	0	5.1
expfrk31	+9.4	+7.7	1028	0	0	6.3
expfrk41	+10.1	+9.8	1285	0	0	6.9
eferk	+9.4	+7.7	512	1	0	6.1
liniger1	+7.3	+3.5	256	1	2	4.9
liniger2	+11.1	+8.5	256	1	2	6.9
liniger3	+11.1	+8.5	256	1	2	6.9

h1 = +.12272₁₀- 1

method	sd1	sd2	fev	jev	lu	time
bw.euler	+1.4	+3.5	513	1	1	14.1
efsirk	+11.4	+11.7	1026	1	2	10.8
expfrk11	+5.8	+3.8	1026	0	0	6.0
expfrk21	+6.3	+8.3	1539	0	0	8.3
expfrk31	+10.9	+9.0	2052	0	0	9.4
expfrk41	+10.8	+9.8	2565	0	0	11.1
eferk	+10.8	+9.0	1024	1	0	12.9
liniger1	+8.5	+4.1	512	1	2	9.8
liniger2	+11.1	+11.2	512	1	2	14.4
liniger3	+11.1	+11.2	512	1	2	13.9

h1 = +.61359₁₀- 2

method	sd1	sd2	fev	jev	lu	time
bw.euler	+1.7	+4.1	1025	1	1	27.5
efsirk	+10.6	+8.9	2048	1	2	22.1
expfrk11	+6.7	+4.4	2048	0	0	11.9
expfrk21	+7.2	+9.1	3072	0	0	16.0
expfrk31	+12.0	+8.9	4096	0	0	18.6
expfrk41	+10.9	+8.9	5120	0	0	22.8
eferk	+12.0	+8.9	2048	1	0	25.1
liniger1	+9.7	+4.7	1024	1	2	20.2
liniger2	+10.5	+8.9	1024	1	2	28.7
liniger3	+10.5	+8.9	1024	1	2	28.2

Opmerkingen:

Uit de resultaten blijkt dat de methoden die het meest geschikt zijn voor het oplossen van stijve vergelijkingen ook bij dit voorbeeld superieur zijn. Voor de expfrk<i><j> formules is het effect van een hogere orde (<i>) duidelijk merkbaar.

Literatuur

- [1] P.W. Hemker, An ALGOL 60 procedure for the solution of stiff differential equations.
MR 128/71, Mathematisch Centrum.
- [2] K. Dekker, Een ALGOL 60 versie van exponentieel aangepaste Runge-Kutta methoden
NR 25/72, Mathematisch Centrum.
- [3] W. Liniger and R.A. Willoughby, Efficient integration methods for stiff systems of ordinary differential equations.
Siam J. Numer. Anal. Vol. 7, no. 1, March 1970.
- [4] R.W. Klopfenstein and C.B. Davis (1970), PECE algorithms for the solution of stiff equations.
Unpublished manuscript.
- [5] C.W. Gear (1968), The automatic integration of stiff ordinary differential equations.
Proceedings of IFIP Congress, Vol. 1, pp. 187-194, North Holland Publishing Comp., Amsterdam.
- [6] W. Liniger and R.A. Willoughby (1967), Efficient numerical integration of stiff systems of ordinary differential equations.
Technical Report RC-1970, IBM Thomas J. Watson Research Center, Yorktown Heights, N.Y..
- [7] A.K. Datta (1967), An evaluation of the approximate inverse algorithm for numerical integration of stiff differential equations.
Technical Report MS/67/84, Imperial Chemical Industries Ltd., Cheshire.

12. Testresultaten van methoden met een stapkeuze-strategie

In dit laatste hoofdstuk zullen we de resultaten geven van een aantal berekeningen, uitgevoerd met een drietal integratie procedures die voorzien zijn van een strategie om de staplengte te variëren.

We hebben gekozen voor de procedures MULTISTEP, EFERK en EFSIRK die reeds in hoofdstuk 11 aan de orde zijn geweest.

Met de testvoorbeelden hebben we ons beperkt tot de stelsels van voorbeeld 3 en voorbeeld 7 uit het vorige hoofdstuk.

We geven nu wat commentaar op de resultaten, die in beeld zijn gebracht in fig. 12.1 (vb.3) en fig. 12.2 (vb.7).

Voorbeeld 3 (zie pag. 259)

Bij MULTISTEP kunnen we een zekere gelaagdheid in de resultaten waarnemen. Iedere "laag" geeft een verzameling van oplossingen, die alle verkregen zijn met dezelfde tolerantie maar met een verschillende minimaal toelaatbare integratie-staplengte (zie tabel 12.1).

Tabel 12.1

Precisietabel voor MULTISTEP, bij verschillende combinaties van minimale staplengte en tolerantie, behorende bij voorbeeld 3 (de precisie wordt gegeven door het aantal juiste decimalen).

tol					
<u>hmin</u>	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
10^{-3}	2.53	3.02	3.50	3.62	-
10^{-4}	2.21	2.69	3.67	4.46	-
10^{-5}	1.77	2.65	3.62	4.50	5.30
10^{-6}	2.10	2.89	-	-	-
10^{-7}	2.09	2.96	3.53	-	-
10^{-8}	2.09	2.96	3.53	4.48	-
10^{-9}	2.09	2.96	3.53	4.48	5.33

Omdat het stapkeuze-mechanisme by EFSIRK alleen effectief is bij niet-lineaire differentiaalvergelijkingen hebben we voor EFSIRK een constante integratie-stap genomen en zodoende een duplicatie gekregen van de resultaten in hoofdstuk 11.

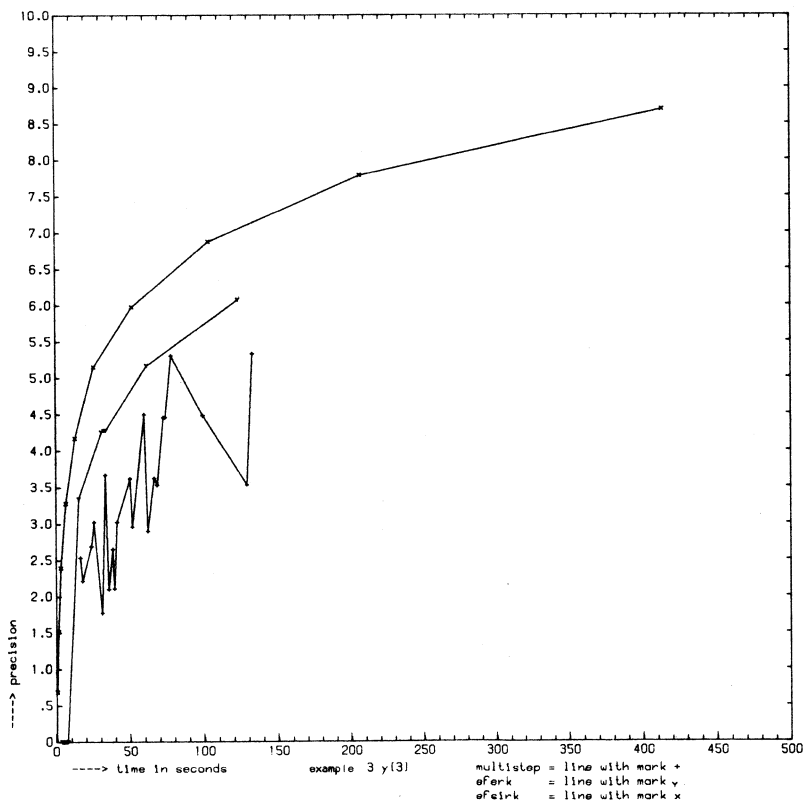
Ook de resultaten verkregen met EFERK geven een gedeeltelijke duplicatie van de resultaten met constante stap uit het vorige hoofdstuk. Voor dit voorbeeld bleek dat een variabele stap bij EFERK geen verbetering brengt in de representatie van de stijve componenten y_1 en y_2 .

We willen tot slot opmerken dat in dit voorbeeld de resultaten verkregen met EFERK de veronderstelling bevestigen dat bij lineaire stelsels het best met een constante stap geïntegreerd kan worden (zie hoofdstuk 6, pag. 150).

Voorbeeld 7 (zie pag. 276)

EFSIRK en EFERK geven ongeveer hetzelfde verloop in de testresultaten te zien. EFERK komt sneller tot de gewenste precisie dan EFSIRK doordat de laatste methode meer berekeningen per stap uitvoert (LU-ontbinding). Opvallend is dat de constante stap strategie uit hoofdstuk 11 voor zowel EFERK als EFSIRK bij voldoende kleine stap gunstiger resultaten te zien geeft.

figuur 12.1



figuur 12.2

