

Multiparametrische exponentieel gefitte Runge-Kutta-methodes

Robin Keppens

Promotor: Prof. dr. Marnix Van Daele

Vakgroep toegepaste wiskunde, informatica en statistiek

Masterproef ingediend tot het behalen van de academische graad van master in de wiskunde

Academiejaar 2020-2021

Multiparametrische exponentieel gefitte Runge-Kutta-methodes

Robin Keppens

Promotor: Prof. dr. Marnix Van Daele

Vakgroep toegepaste wiskunde, informatica en statistiek

Masterproef ingediend tot het behalen van de academische graad van master in de wiskunde

Academiejaar 2020-2021

Voorwoord

Gedurende mijn opleiding ben ik in contact gekomen met verschillende takken van de wiskunde, waarvan de schoonheid telkens door fervente professoren en assistenten verkondigd werd, vaak met succes. Toch is mijn interesse naarmate mijn laatste masterjaar naderde steeds meer verschoven naar de toegepaste wiskunde, en meer in het bijzonder naar de toepassingen waarbij ik mijn voorliefde voor programmeren kon gebruiken. In dat opzicht leek een masterproef rond numerieke methodes mij dus perfect. Toen ik vervolgens in het vak *Capita selecta in de numerieke wiskunde* exponentieel gefitte methodes leerde kennen en er mee moest experimenteren, was ik helemaal verkocht. Ik ben dan ook zeer dankbaar dat ik de kans heb gekregen om verder te werken met wat daar aan bod gekomen is.

Het was geen eenvoudige zaak om in tijden van corona een masterproef te schrijven: geen afwisselende contacten met vrienden op de campus, enkel online overlegmomenten met de promotor, amper activiteiten ter ontspanning mogelijk, ... Maar ik hoop toch dat ik er in ben geslaagd om een interessante masterproef te schrijven waar ik met trots op kan terugkijken. Ik was in elk geval nooit zo ver geraakt zonder de morele steun van Lara, bedankt om er voor mij te zijn. Ook bedankt aan mijn ouders voor het vertrouwen, ik weet dat dat niet evident was. Tot slot wil ik professor Van Daele bedanken voor de persoonlijke begeleiding gedurende het afgelopen jaar, zonder de vele tips en het samen overleggen zou deze masterproef niet gelukt zijn.

De auteur geeft de toelating deze masterproef voor consultatie beschikbaar te stellen en delen van de masterproef te kopiëren voor persoonlijk gebruik. Elk ander gebruik valt onder de beperkingen van het auteursrecht, in het bijzonder met betrekking tot de verplichting de bron uitdrukkelijk te vermelden bij het aanhalen van resultaten uit deze masterproef.

A handwritten signature in dark ink, appearing to read 'Y. Keppens', is enclosed within a large, loopy oval shape. A horizontal line extends from the right side of the oval.

Dendermonde, 1 juni 2021

Inhoudsopgave

| | | |
|----------|--|-----------|
| 1 | Introductie | 1 |
| 2 | Basisconcepten van RK-methodes | 3 |
| 2.1 | Definitie | 3 |
| 2.2 | Bepaling van de coëfficiënten | 3 |
| 2.3 | Symplecticiteit | 5 |
| 2.4 | Symmetrie | 5 |
| 2.5 | Foutterm | 5 |
| 2.5.1 | Standaard RK-methodes | 5 |
| 2.5.2 | EFRK-methodes | 6 |
| 2.5.3 | Gemodificeerde EFRK-methodes | 7 |
| 2.5.4 | Herschikking | 9 |
| 2.6 | Stabiliteit | 11 |
| 3 | Tweetrapsmethode | 12 |
| 3.1 | De methode | 12 |
| 3.2 | Coëfficiënten | 15 |
| 3.3 | Foutterm | 17 |
| 3.3.1 | Bomen van orde één | 17 |
| 3.3.2 | Bomen van orde twee | 18 |
| 3.3.3 | Bomen van orde drie | 18 |
| 3.3.4 | Bomen van orde vier | 18 |
| 3.3.5 | Bomen van orde vijf | 18 |
| 3.3.6 | Spookbomen | 19 |
| 3.3.7 | Interpretatie | 19 |
| 3.4 | Stabiliteit | 19 |
| 3.5 | Speciale gevallen | 21 |
| 4 | Drietrapsmethode | 23 |
| 4.1 | De methode | 23 |
| 4.2 | Coëfficiënten | 26 |
| 4.3 | Foutterm | 28 |
| 4.4 | Stabiliteit | 31 |
| 4.5 | Speciale gevallen | 33 |
| 5 | Viertrapsmethode | 35 |
| 5.1 | De methode | 35 |
| 5.2 | Coëfficiënten | 37 |
| 6 | Numerieke resultaten | 38 |
| 6.1 | Het geperturbeerd keplerprobleem | 38 |
| 6.1.1 | Tweetrapsmethode | 39 |
| 6.1.2 | Drietrapsmethode | 39 |
| 6.1.3 | Viertrapsmethode | 40 |

| | | |
|----------|--|-----------|
| 6.2 | Een autonoom en lineair probleem | 41 |
| 6.2.1 | Tweetrapsmethode | 41 |
| 6.2.2 | Drietrapsmethode | 43 |
| 6.2.3 | Viertrapsmethode | 44 |
| 6.3 | Het keplerprobleem | 45 |
| 6.3.1 | Tweetrapsmethode | 46 |
| 6.3.2 | Drietrapsmethode | 47 |
| 6.4 | Het eulerprobleem | 48 |
| 6.4.1 | Tweetrapsmethode | 48 |
| 6.4.2 | Drietrapsmethode | 49 |
| 7 | Besluit | 51 |
| | Appendix | 52 |
| A | English summary | 52 |
| A.1 | Two-stage method | 52 |
| A.2 | Three-stage method | 52 |
| A.3 | Four-stage method | 53 |
| A.4 | Conclusion | 53 |
| | Referenties | 54 |
| | Lijst van programma's | 56 |

1 Introductie

Het oplossen van (stelsels van) differentiaalvergelijkingen komt voor in diverse takken van de toegepaste wetenschappen: mechanica, astrofysica, moleculaire dynamica, ... In veel van deze gevallen vertoont de oplossing een oscillatorisch of exponentieel karakter en bijgevolg is er in de afgelopen decennia veel onderzoek gedaan naar het bepalen van numerieke oplossingen die dit karakter behouden. Een belangrijk resultaat hiervan is het ontstaan van exponentieel gefitte methodes (verder gebruikt als EF, naar *exponentially fitted*): waar men traditioneel eiste dat de methodes exact zijn voor een collectie veeltermfuncties, werd de focus nu verlegd naar functies van de vorm $y(x) = \sum_{i=1}^I f_i(x) \exp(\mu_i x)$, waarbij $f_i(x)$ polynomen zijn van een lage graad. De parameters μ_i kunnen enerzijds reëel zijn en geven dan aanleiding tot exponentiële functies, en anderzijds zuiver imaginair, waardoor trigonometrische functies verkregen worden. Voor de concrete waarden, die staan voor de werkelijke frequenties in het probleem domein, bestaan soms exacte waarden of goeie schattingen, maar vaak zijn ze het onderwerp van een onderzoek op zich.

Een van de manieren om een exponentieel gefitte methode te ontwikkelen is om te vertrekken van een klassieke methode met parameters en voor een verzameling functies S , de zogenaamde *fitting space*, op te leggen dat ze exact worden opgelost. Hieruit worden dan de waarden voor de parameters verkregen. Een goeie keuze voor S bestaat meestal uit de unie van enerzijds de veeltermfuncties tot een zekere graad k en anderzijds een verzameling exponentiële functies met verschillende frequenties μ_i . De klassieke methodes waarmee in deze masterproef wordt vertrokken zijn de Runge-Kutta-methodes (RK-methodes) en de exponentieel gefitte uitbreidingen worden verkort EFRK-methodes genoemd. In dit onderzoeksdomein zijn reeds verschillende resultaten bekend. Zo hebben Van den Berghe et al. (zie [16]) het geval bekeken waarbij de μ_i allemaal gelijk gesteld worden, wat resulteert in

$$S_{k,l} = \{1, x, \dots, x^k\} \cup \{\exp(\pm\mu x), x \exp(\pm\mu x), \dots, x^l \exp(\pm\mu x)\}.$$

Anderzijds onderzochten Calvo et al. (zie [2]-[3]) het geval van de veelvouden van een vaste frequentie μ , dit geeft

$$S_{k,l} = \{1, x, \dots, x^k\} \cup \{\exp(\pm\mu x), \exp(\pm 2\mu x), \dots, \exp(\pm l\mu x)\}.$$

Elk van deze papers bleven echter beperkt tot één frequentie μ en variaties daarvan. Nochtans maken beide gevallen deel uit van een algemenere verzameling

$$S_{k,l} = \{1, x, \dots, x^k\} \cup \{\exp(\pm\mu_1 x), \exp(\pm\mu_2 x), \dots, \exp(\pm\mu_l x)\}.$$

Het is duidelijk dat deze ongerelateerde frequenties enerzijds mogelijkheden openen om bepaalde problemen nauwkeuriger op te lossen. Anderzijds wordt een van de moeilijkheden uit de resultaten van onder andere Calvo en Van den Berghe, namelijk de bepaling van een goeie frequentie μ , nog lastiger. De bedoeling van deze masterproef is het bekijken van deze moeilijkheid voor enkele relatief eenvoudige gevallen: tweetraps, drietraps en viertraps EFRK-methodes. In hoofdstuk 2 wordt de wiskundige basis gegeven die nodig is om alle aspecten van deze methodes te behandelen, samen met de definiëring van enkele bijkomende eigenschappen die zullen worden opgelegd omdat ze vaak voordelig zijn. In hoofdstukken 3 en 4 worden dan respectievelijk het tweetraps- en het drietrapsgeval afgeleid, waarbij de

coëfficiënten telkens onder de loep genomen worden en de foutterm berekend wordt, die belangrijke informatie kan geven. Telkens worden ook de stabiliteitseigenschappen bekeken en wordt de link gelegd met de reeds bestaande methodes. Vervolgens wordt in hoofdstuk 5 het viertrapsgeval beperkter behandeld, met enkel de afleiding en een bespreking van de coëfficiënten. In hoofdstuk 6 wordt uiteindelijk een poging gedaan om het eerder vermelde probleem aan te pakken: voor enkele bekende problemen worden de numerieke oplossingen bekeken voor verschillende configuraties van de parameters. Tot slot worden in hoofdstuk 7 de belangrijkste resultaten opgelijst.

2 Basisconcepten van RK-methodes

2.1 Definitie

Een algemene s -traps RK-methode kan als volgt gekarakteriseerd worden:

$$\begin{cases} y_{n+1} &= y_n + h \sum_{i=1}^s b_i f(x_n + c_i h, Y_i) \\ Y_i &= y_n + h \sum_{j=1}^s a_{ij} f(x_n + c_j h, Y_j) \quad i = 1, \dots, s \end{cases} \quad (1)$$

en wordt vaak verkort voorgesteld door het corresponderende Butcher-tableau:

$$\begin{array}{c|c} c & A \\ \hline & b \end{array} = \begin{array}{c|ccc} c_1 & a_{11} & \dots & a_{1s} \\ \vdots & & \ddots & \\ c_s & a_{s1} & \dots & a_{ss} \\ \hline & b_1 & \dots & b_s \end{array} \quad (2)$$

Om tot een ruimere verzameling functies te komen die exact worden geïntegreerd, zijn meer vrijheidsgraden nodig. Hiertoe werden de gemodificeerde RK-methodes geïntroduceerd:

$$\begin{cases} y_{n+1} &= y_n + h \sum_{i=1}^s b_i f(x_n + c_i h, Y_i) \\ Y_i &= \gamma_i y_n + h \sum_{j=1}^s a_{ij} f(x_n + c_j h, Y_j) \quad i = 1, \dots, s \end{cases} \quad (3)$$

met als bijbehorende tableau

$$\begin{array}{c|c|c} c & \Gamma & A \\ \hline & & b \end{array} = \begin{array}{c|cc|ccc} c_1 & \gamma_1 & a_{11} & \dots & a_{1s} \\ \vdots & \vdots & & \ddots & \\ c_s & \gamma_s & a_{s1} & \dots & a_{ss} \\ \hline & & b_1 & \dots & b_s \end{array} \quad (4)$$

Een uitgebreidere behandeling van RK-methodes met alternatieve definities, basiseigenschappen en voorbeelden kan gevonden worden in [12].

Opmerking 1. *In theorie maken methodes waarbij de externe vergelijking gegeven wordt door $y_{n+1} = \gamma_0 y_n + h \sum_{i=1}^s b_i f(x_n + c_i h, Y_i)$ ook deel uit van de gemodificeerde RK-methodes. Methodes waar $\gamma_0 \neq 1$ voldoen echter niet aan de consistentievoorwaarde dat constante functies exact worden opgelost en worden bijgevolg niet beschouwd.*

2.2 Bepaling van de coëfficiënten

Om tot concrete waarden te komen voor de coëfficiënten zal hier gewerkt worden met de eerder vermelde fitting space S . De voorwaarden die hieruit volgen voor $f(x) \in S$ kunnen

worden uitgedrukt met lineaire functionalen:

$$\begin{aligned}\mathcal{L}_i[f(x), \mathbf{a}] &= f(x_n + c_i h) - \gamma_i f(x_n) - h \sum_{j=1}^s a_{ij} f'(x_n + c_j h) = 0, \quad 1 \leq i \leq s, \\ \mathcal{L}[f(x), \mathbf{b}] &= f(x_n + h) - f(x_n) - h \sum_{i=1}^s b_i f'(x_n + c_i h) = 0.\end{aligned}\tag{5}$$

Bij RK-methoden bestaat ook de mogelijkheid om twee verschillende verzamelingen S_{int} en S_{ext} te definiëren, waaraan respectievelijk de interne trappen met $\mathcal{L}_i[f(x), \mathbf{a}]$ en de externe trappen met $\mathcal{L}[f(x), \mathbf{b}]$ voldoen. Bij EFRK-methoden in het bijzonder wordt vaak met de volgende algemene vorm van de fitting space gewerkt:

$$S_{k,l} = \{1, x, \dots, x^k\} \cup \{\exp(\pm\mu_1 x), \exp(\pm\mu_2 x), \dots, \exp(\pm\mu_l x)\}.$$

De exacte waarden van k en l hangen nu af van het aantal vrijheidsgraden van het systeem: voor een s -traps RK-methode zijn er initieel $s + 1$ vrije keuzes in de interne trappen en $2s$ in de externe trap. Verdere restricties op het gedrag van de methode, zoals symmetrie en symplecticiteit (zie verder), verlagen uiteraard het aantal mogelijkheden, maar geven ook bepaalde functies uit de fitting space ‘cadeau’. De uiteindelijke bepaling van de coëfficiënten in de methodes die verder zijn opgelijst is dan ook vaak het gevolg van het blijven opleggen van extra voorwaarden van de vorm (5), tot wanneer alle vrije parameters bepaald zijn, om vervolgens te testen of er eventueel andere functies ook exact worden geïntegreerd.

In de fitting space komen exponentiële functies voor die zowel een reëel als een complex argument kunnen hebben. Bijgevolg zal deze afhankelijkheid zich ook laten zien in de uitdrukkingen van de coëfficiënten: bij reële frequenties kunnen ze herschreven worden in termen van hyperbolische functies, bij imaginaire frequenties in termen van goniometrische functies. Op zich is dit geen probleem, en bovendien zijn beide soorten functies voor beide soorten frequenties gedefinieerd. Het is echter mooier en vaak eenvoudiger om te werken met de gekwadraterde versie. Elk van de coëfficiënten zal uiteindelijk in functie zijn van zekere $z_i = \mu_i h$, en dus na de overgang in functie van $Z_i = z_i^2$. Dit heeft enerzijds het voordeel dat de afhankelijke parameter reëel is en dus makkelijker weer te geven in plots, en anderzijds blijken de coëfficiënten vaak ook makkelijk te herschrijven met behulp van de zogenaamde functies van Ixaru (zie [7]), recursief gedefinieerd door:

$$\begin{aligned}\eta_{-1}(Z) &= \begin{cases} \cos(\sqrt{|Z|}) & Z \leq 0 \\ \cosh(\sqrt{Z}) & Z > 0 \end{cases} & \eta_0(Z) &= \begin{cases} \frac{\sin(\sqrt{|Z|})}{\sqrt{|Z|}} & Z < 0 \\ 1 & Z = 0 \\ \frac{\sinh(\sqrt{Z})}{\sqrt{Z}} & Z > 0 \end{cases} \\ \eta_m(Z) &= \begin{cases} \frac{\eta_{m-2}(Z) - (2m-1)\eta_{m-1}(Z)}{Z} & Z \neq 0 \\ \frac{1}{(2m+1)!!} & Z = 0 \end{cases} & m \in \mathbb{N}_0\end{aligned}\tag{6}$$

In de rest van deze masterproef zal met Z_i steeds z_i^2 bedoeld worden.

2.3 Symplecticiteit

In veel gevallen maakt de oscillerende oplossing die gefit wordt deel uit van een hamiltoniaans systeem. Een uitermate geschikte subklasse van integrators voor deze systemen zijn de symplectische methoden, die de symplecticiteit van de stroming van het systeem behouden en eveneens de hamiltoniaan tot op de orde van de methode benaderen (zie onder meer [13]). Het volgende resultaat is bewezen omtrent de voorwaarden voor symplecticiteit bij gemodificeerde RK-methoden (zie [15], die voor het bewijs steunt op [11]):

Stelling 2. *Een gemodificeerde s-traps RK-methode is symplectisch als en slechts als de coëfficiënten aan de volgende vergelijkingen voldoen:*

$$\frac{b_j}{\gamma_j}a_{ji} + \frac{b_i}{\gamma_i}a_{ij} - b_i b_j = 0, \quad 1 \leq i, j \leq s.$$

Een bijkomend voordeel is dat, indien aan de voorwaarden van stelling 2 is voldaan, de methode ook kwadratische invarianten behoudt, bewezen in [4].

2.4 Symmetrie

Het is vaak ook nuttig dat, wanneer de methode achterwaarts wordt toegepast, men terug bij het originele beginpunt uitkomt. Over lange intervallen hebben dergelijke methodes immers een beter gedrag dan andere, bewezen door Hairer et. al in [6]. Voor een correcte definiëring van het begrip symmetrie wordt gesteund op de volgende definitie:

Definitie 3. *De toegevoegde methode ψ_h^* van een numerieke methode ψ_h met stapgrootte h is de inverse afbeelding van de originele methode met stapgrootte $-h$, of dus $\psi_h^* = \psi_{-h}^{-1}$.*

Een methode wordt nu symmetrisch genoemd indien hij gelijk is aan zijn toegevoegde methode: $\psi_h^* = \psi_h$. De voorwaarden voor het symmetrisch zijn van een gemodificeerde RK-methode worden in de volgende stelling uitgedrukt (zie [2]).

Stelling 4. *Een gemodificeerde s-traps RK-methode is symmetrisch als en slechts als de coëfficiënten aan de volgende vergelijkingen voldoen:*

$$\begin{cases} c_i &= 1 - c_{s+1-i} \\ b_i &= b_{s+1-i} \\ \gamma_i &= \gamma_{s+1-i} \\ a_{i,j} &= \gamma_i b_j - a_{s+1-i, s+1-j} \end{cases} \quad 1 \leq i, j \leq s.$$

2.5 Foutterm

2.5.1 Standaard RK-methodes

De foutterm van een standaard RK-methode kan worden berekend met behulp van de theorie van Butcher:

$$y(x_{n+1}) - y_{n+1} = \sum_{q=1}^{\infty} \frac{h^q}{q!} \sum_{\rho(t)=q} \alpha(t)[1 - \gamma(t)\psi(t)]F(t). \quad (7)$$

Hierbij loopt de tweede sommatie over alle gewortelde bomen \mathcal{T} , waarvoor een numerieke behandeling terug te vinden is in [12]. De functies ρ en γ omvatten eigenschappen van de bomen.

Definitie 5. De orde $r(t)$, de symmetrie $\sigma(t)$ en de dichtheid $\gamma(t)$ van een boom $t \in \mathcal{T}$ worden recursief gedefinieerd door:

$$\begin{aligned}\rho(\tau) &= \sigma(\tau) = \gamma(\tau) = 1, \\ \rho([t_1^{n_1} \dots t_k^{n_k}]) &= 1 + n_1\rho(t_1) + \dots + n_k\rho(t_k), \\ \sigma([t_1^{n_1} \dots t_k^{n_k}]) &= n_1! \dots n_k! \sigma(t_1)^{n_1} \dots \sigma(t_k)^{n_k}, \\ \gamma([t_1^{n_1} \dots t_k^{n_k}]) &= \rho([t_1^{n_1} \dots t_k^{n_k}]) \gamma(t_1)^{n_1} \dots \gamma(t_k)^{n_k}.\end{aligned}$$

Met behulp hiervan kan de derde voorkomende functie α bepaald worden, die staat voor het aantal essentieel verschillende manieren om de knopen van een boom te labelen.

$$\alpha(t) = \frac{\rho(t)!}{\sigma(t)\gamma(t)}$$

De eigenlijke ordevoorwaarden zitten dan vervat in de factor $[1 - \gamma(t)\psi(t)]$: als geldt dat $\psi(t) = \gamma(t)^{-1}$ voor alle bomen t van orde 1 tot en met q , dan is de leidende foutterm van orde $\mathcal{O}(h^{q+1})$ en is de methode dus van orde q . De functie ψ is afhankelijk van de coëfficiënten van de RK-methode.

Definitie 6. Voor $t \in \mathcal{T}$ wordt $\psi(t)$ gedefinieerd als:

$$\begin{aligned}\psi_i(\tau) &= \sum_{j=1}^s a_{ij}, & \psi_i([t_1 \dots t_k]) &= \sum_{j=1}^s a_{ij} \psi_j(t_1) \dots \psi_j(t_k), & i &= 1, \dots, s+1, \\ \psi(t) &= \psi_{s+1}(t).\end{aligned}$$

Hierbij is $a_{s+1,j} = b_j$.

De factor $F(t)$ tot slot is de elementaire differentiaal corresponderend met t , geëvalueerd in $y(x_n)$. De afleiding van dit resultaat en een ruimere bespreking van deze begrippen kan gevonden worden in [8] en [12].

2.5.2 EFRK-methodes

Bij EFRK-methodes zijn de coëfficiënten functies van $z_i = h\mu_i$. Aangezien de foutterm een reeksontwikkeling in h is, dienen deze coëfficiënten (die opduiken in de factoren $\psi(t)$) eveneens in reeks ontwikkeld te worden, waarna de termen met dezelfde orde van h pas kunnen samengenomen worden om de correcte orde af te lezen. De voorwaarde opdat een methode van orde q is, wordt dus als volgt samengevat:

Lemma 7. Een EFRK-methode is van orde q als en slechts als $\psi(t) = \gamma(t)^{-1} + \mathcal{O}(h^{q+1-\rho(t)})$ voor alle bomen t tot en met orde q .

Er moet dus aan meer voorwaarden voldaan zijn dan in het klassieke geval, maar omdat de functies van de gebruikte fitting space symmetrisch zijn (steeds μ_i en $-\mu_i$), kan een belangrijke vereenvoudiging gedaan worden. Er geldt immers dat de reeksontwikkelingen van de coëfficiënten even zijn, of dus:

$$\left. \frac{\partial^{2k-1} a_{ij}}{\partial h^{2k-1}} \right|_{h=0} = \left. \frac{\partial^{2k-1} b_i}{\partial h^{2k-1}} \right|_{h=0} = 0, \quad k \in \mathbb{N}_0, \quad 1 \leq i, j \leq s.$$

Hieruit volgt dan dat voor alle bomen t de reeksontwikkeling van $\psi(t)$, een product van a_{ij} 's en b_i 's, ook even is:

$$\left. \frac{\partial^{2k-1} \psi(t)}{\partial h^{2k-1}} \right|_{h=0} = 0, \quad k \in \mathbb{N}_0.$$

Bijgevolg valt de helft van de ordetermen in h weg.

2.5.3 Gemodificeerde EFRK-methodes

Voor een gemodificeerde EFRK-methode klopt uitdrukking (7) voor de foutterm niet meer. Om dit probleem te illustreren kan gekeken worden naar een eenvoudige expliciete twee-trapsmethode:

$$\begin{array}{c|c|cc} 0 & 1 & 0 & 0 \\ 1 & \gamma & a & 0 \\ \hline & & b & b \end{array}$$

Nu is ook γ een functie van z_i en dus van h , en de reeksontwikkeling bevat opnieuw alleen even termen. Stel $\gamma = 1 + \delta$, dan is $\delta = h^2 \delta_0 + \mathcal{O}(h^4)$. Ook a en b kunnen als reeks worden geschreven: $a = a_0 + h^2 a_2 + \mathcal{O}(h^4)$ en $b = b_0 + h^2 b_2 + \mathcal{O}(h^4)$. Toegepast op het autonome probleem $y' = f(y)$ levert de Taylorontwikkeling van de numerieke oplossing in $h = 0$ tot op derde orde (met $y_n = y$ en $f(y_n) = f$):

$$\begin{aligned} y_{n+1} &= y + hb[f + f(\gamma y + haf)] \\ &= y + h(b_0 + h^2 b_2)[f + f + ha_0 f' f + \frac{h^2}{2} a_0^2 f'' f^2 + h^2 a_2 f' f + h^2 \delta_0 f' y] + \mathcal{O}(h^4) \\ &= y + 2hb_0 f + h^2 a_0 b_0 f' f + 2h^3 b_2 f + \frac{h^3}{2} a_0^2 b_0 f'' f^2 + h^3 a_2 b_0 f' f + \mathbf{h^3 b_0 \delta_0 f' y} + \mathcal{O}(h^4) \end{aligned}$$

De vetgedrukte term, die in het klassieke geval niet voorkomt (daar is $\delta = 0$), is een zogenaamde *spookterm*. Dergelijke termen duiken vanaf derde orde op en zijn nu ook geassocieerd aan elementaire differentiaal, zogenaamde spookdifferentialen. Ze dienen bijgevolg ook opgenomen te worden in de foutterm met behulp van spookbomen:






Definitie 8. De verzameling spookbomen \mathcal{G} wordt als volgt recursief gedefinieerd:

- $\tau_g = \circ \in \mathcal{G}$ (de boom geassocieerd aan y),
- $t_1, \dots, t_k \in \mathcal{G} \cup \mathcal{T}, \exists i : t_i \in \mathcal{G} \Rightarrow t = [t_1 \dots t_k] \in \mathcal{G}$.

De gecorrigeerde versie voor de foutterm wordt dan (zie [9]):

$$y(x_{n+1}) - y_{n+1} = \sum_{q=1}^{\infty} \frac{h^q}{q!} \sum_{\rho(t)=q} \alpha(t)[1 - \gamma(t)\psi(t)]F(t) - \sum_{t \in \mathcal{G}, \rho(t) \geq 2} \frac{h^{\rho'(t)}}{\sigma(t)} \psi(t)F(t). \quad (8)$$

Hiervoor dient definitie 6 voor ψ uitgebreid te worden naar spookbomen met behulp van de betrekking $\psi_i(\tau_g) = \delta_i$, terwijl voor de andere functies uit definitie 5 de boom τ_g hetzelfde wordt behandeld als τ . De nieuwe functie ρ' wordt gedefinieerd als $\rho'(t) = \rho(t) - \rho_g(t)$, waarbij de spookorde $\rho_g(t)$ het aantal bladeren in een boom t teruggeeft corresponderend met de spookboom τ_g . In tabel 1 worden een aantal spookbomen samen met de corresponderende spookdifferentialen opgelijst.

| Spookboom | Spookdifferential | $\rho(t)$ | $\rho'(t)$ |
|---|-------------------|-----------|------------|
|  | y | 1 | 0 |
|  | $\{y\}$ | 2 | 1 |
|  | $\{fy\}$ | 3 | 1 |
|  | $\{y^2\}$ | 3 | 2 |
|  | $\{2y\}_2$ | 3 | 1 |

Tabel 1: Alle spookbomen t met $\rho(t) \leq 3$.

De voorwaarde uit lemma 7 om een methode van orde q te hebben is nu niet langer voldoende, er dient nu ook te gelden dat $\psi(t) = 0$ voor alle spookbomen t waarvoor $\rho'(t) \leq q$. Ook hier moet rekening gehouden worden met het feit dat de coëfficiënten in de spooktermen zelf in functie van h zijn en er dus herschikkingen nodig zijn. Uiteindelijk wordt het volgende resultaat bekomen:

Lemma 9. *Een gemodificeerde EFRK-methode is van orde q als en slechts als:*

- $\psi(t) = \gamma(t)^{-1} + \mathcal{O}(h^{q+1-\rho(t)})$ voor alle bomen t tot en met orde q .
- $\psi(t) = \mathcal{O}(h^{q+1-\rho'(t)})$ voor alle spookbomen t waarvoor $\rho'(t) \leq q$.

De sommatie over de spookbomen kan hier inzichtelijker gemaakt worden. Zoals eerder vermeld is voor de δ_i de reeksontwikkeling in h minstens van orde $\mathcal{O}(h^2)$. Elk voorkomen van een δ_i in de uitdrukking $\psi(t)$, of analoog het aantal voorkomens $\rho_g(t)$ van de spookboom τ_g als deelboom in een boom t (uit de definitie van ψ), verhoogt bijgevolg de orde van h in de term bij t met minstens twee. De totale orde van h wordt dan minstens $\rho'(t) + 2\rho_g(t) = \rho(t) + \rho_g(t)$ voor een zekere boom t . De voorwaarde $\psi(t) = \mathcal{O}(h^{q+1-\rho'(t)})$ moet dus enkel gecontroleerd worden voor bomen t met $\rho(t) + \rho_g(t) \leq q$.

Voorbeeld 10. *Neem de spookboom $t = [\tau_g]$. De bijbehorende spookterm wordt gegeven door:*

$$-\frac{h^{\rho'(t)}}{\sigma(t)}\psi(t)F(t) = -h \sum_{j=1}^s b_j \delta_j \{y\}.$$

In eerste instantie lijkt het dus alsof deze term ook een bijdrage kan leveren aan de eerste-ordeterm van de reeksontwikkeling van de fout. De som bevat echter eenmaal δ_j , omdat τ_g één keer voorkomt als deelboom en in de definitie van ψ elke τ_g wordt omgezet in δ_j . Omdat $\delta_j = \mathcal{O}(h^2)$, zal er uiteindelijk minstens h^3 voorop staan, dus voor methodes van orde twee is $\psi(t) = \mathcal{O}(h^{q+1-\rho'(t)})$ automatisch voldaan, ook al is $\rho'(t) = 1 < q = 2$.

2.5.4 Herschikking

Als een methode dan effectief van orde q is, dan zal de leidende term van de fout (ook lte genoemd, naar *leading term error*) van de volgende vorm zijn:

$$lte = h^{q+1} \sum_{t \in \mathcal{T} \cup \mathcal{G}} c_t F(t).$$

De elementaire differentiaal $F(t)$ kunnen nu per orde verder gegroepeerd worden tot gewone afgeleiden van y . In het klassieke geval kunnen voor de leidende term in de afknottingsfout de elementaire differentiaal herschreven worden in termen van totale afgeleiden van aflopende ordes $q+1$ tot en met $r+1$ (zie [1]), waarbij r de traporde van de methode is.

Definitie 11. *De traporde van een s -traps RK-methode is de hoogst mogelijke r waarvoor geldt dat*

$$\sum_{j=1}^s a_{ij} c_j^{k-1} = \frac{c_i^k}{k}, \quad i = 1, \dots, s, \quad k = 1, \dots, r.$$

Hiermee kan gewerkt worden naar een meer inzichtelijke vorm van de leidende foutterm:

$$lte = h^{q+1} \sum_{i=r+1}^{q+1} c_i \times \text{‘afgeleiden van } f \text{’} \times y^{(j)}.$$

De ‘afgeleiden van f ’ zijn dan jacobianen en andere Fréchetafgeleiden (zie [12] voor een definitie van dit begrip).

Voorbeeld 12. *Beschouw de trapeziumregel, een tweetrapsmethode gegeven door het volgende tableau:*

| | | | |
|---|---|---------------|---------------|
| 0 | 1 | 0 | 0 |
| 1 | 1 | $\frac{1}{2}$ | $\frac{1}{2}$ |
| | | $\frac{1}{2}$ | $\frac{1}{2}$ |

Dit is methode van orde 2 en traporde 2 en de lte kan dus herschreven worden als een totale afgeleide van orde 3:

$$\begin{aligned}
lte &= \frac{h^3}{3!} \sum_{\rho(t)=3} \alpha(t)[1 - \gamma(t)\psi(t)]F(t) \\
&= \frac{h^3}{3!} ([1 - 3\psi([\tau^2])]\{f^2\} + [1 - 6\psi([{}_2\tau]_2)]\{{}_2f\}_2) \\
&= \frac{h^3}{3!} \left([1 - 3 \sum_i b_i c_i^2]\{f^2\} + [1 - 6 \sum_{i,j} b_i a_{ij} c_j]\{{}_2f\}_2 \right) \\
&= \frac{h^3}{3!} \left(-\frac{1}{2}\{f^2\} - \frac{1}{2}\{{}_2f\}_2 \right) \\
&= -\frac{h^3}{12}(\{f^2\} + \{{}_2f\}_2) \\
&= -\frac{h^3}{12}y^{(3)}.
\end{aligned}$$

Voor het exponentieel gefitte geval is het herschrijven in gewone afgeleiden niet altijd meer mogelijk. De definitie van de traporde houdt geen rekening met het feit dat de coëfficiënten ook functies van h zijn, dus de gelijkheid in definitie 11 zal maar gelden tot op een bepaalde orde van h . Bovendien leveren bomen van lagere orde vaak ook nog een bijdrage aan de leidende foutterm. Het lukt echter vaak toch om een meer inzichtelijke vorm te verkrijgen.

Voorbeeld 13. De exponentieel gefitte versie van de trapeziumregel op $S_{ext} = \{1, \exp(\mu_1 x)\}$ wordt gegeven door het tableau

$$\begin{array}{c|c|c|c}
0 & 1 & 0 & 0 \\
1 & 1 & b & b \\
\hline
& & b & b
\end{array}$$

met $b = \frac{\sinh(\frac{z_1}{2})}{z_1 \cosh(\frac{z_1}{2})}$. De methode blijft van orde twee, maar de leidende foutterm wordt nu aangevuld met een bijdrage van de boom van orde één.

$$\begin{aligned}
lte &= -\frac{h^3}{12}y^{(3)} + \frac{h}{1!} \sum_{\rho(t)=3} \alpha(t)[1 - \gamma(t)\psi(t)]F(t) \\
&= -\frac{h^3}{12}y^{(3)} + h\alpha(\tau)[1 - \gamma(\tau)\psi(\tau)]F(\tau) \\
&= -\frac{h^3}{12}y^{(3)} + h[1 - \sum_{i=1}^s b_i]f \\
&= -\frac{h^3}{12}y^{(3)} + h[1 - (1 - \frac{1}{12}z_1^2 + \mathcal{O}(h^4))]f \\
&= -\frac{h^3}{12}y^{(3)} + \frac{h}{12}z_1^2 f \\
&= \frac{h^3}{12}(\mu_1^2 y^{(1)} - y^{(3)})
\end{aligned}$$

De lte is nu ook afhankelijk van de eerste orde afgeleide van y , maar er kan wel duidelijk worden afgelezen dat de keuze $\mu_1^2 = \frac{y^{(3)}}{y^{(1)}}$ de leidende foutterm (en ook de andere termen) nul maakt, of dus dat de methode inderdaad exact is voor $\exp(\mu_1 x)$.

2.6 Stabiliteit

Nog een wenselijke eigenschap voor numerieke methoden is dat hun oplossingen hetzelfde gedrag vertonen op oneindig als de exacte oplossing: convergeert de numerieke oplossing naar nul als de exacte oplossing dat doet? Om op deze vraag te antwoorden wordt gebruik gemaakt van een testvergelijking:

$$y'(x) = \omega y(x), \quad \omega \in \mathbb{C}^-. \quad (9)$$

De exacte oplossing van deze testvergelijking is $y = \exp(\omega x) + c$ en convergeert inderdaad slechts naar 0 voor $x \rightarrow \infty$ op voorwaarde dat ω in het complexe linkerhalfvlak ligt. Om te kijken wat er in het numerieke geval gebeurt, wordt gekeken naar het gedrag van de zogenaamde stabiliteitsveelterm $R(z)$. Hier is $z = \omega h$ een complex getal in functie van de frequentie ω van de testvergelijking en de stapgrootte h van de methode. Voor een gemodificeerde RK-methode wordt de stabiliteitsveelterm gegeven door de volgende formule:

$$R(z) = 1 + zb^\top (I - zA)^{-1} \Gamma. \quad (10)$$

De coëfficiënten, bevat in A , b en Γ , zijn bij EFRK-methodes afhankelijk van de frequenties waarop gefit wordt en die afhankelijkheid wordt dan ook doorgetrokken naar $R(z)$ (impliciet staat er dus $R(z, \mu_1, \dots, \mu_l)$). Voor deze veelterm geldt nu dat

$$y_{n+1} = R(z)y_n = R(z)^{n+1}y_0.$$

Hiermee is het duidelijk dat de convergentie op oneindig van de numerieke oplossing pas gewaarborgd is indien $|R(z)| < 1$. Er moet dus gezocht worden naar het gebied van absolute stabiliteit:

$$\mathcal{R}_A = \{z \mid |R(z)| < 1\}. \quad (11)$$

Is \mathbb{C}^- bevat in \mathcal{R}_A , dan is het gedrag op oneindig dus overeenkomstig en spreekt men van een absoluut stabiele of A-stabiele methode. Indien de eigenschap $\mathbb{C}^- = \mathcal{R}_A$ geldt, dan wordt de methode exact-A-stabiel genoemd.

3 Tweekrapsmethode

3.1 De methode

De standaard tweekraps gemodificeerde RK-methode ziet er als volgt uit:

$$\begin{array}{c|c|cc} c_1 & \gamma_1 & a_{11} & a_{12} \\ c_2 & \gamma_2 & a_{21} & a_{22} \\ \hline & & b_1 & b_2 \end{array}$$

De symmetrievorwaarden worden hier

$$\left\{ \begin{array}{l} c_1 = 1 - c_2 \\ \gamma_1 = \gamma_2 = \gamma \\ b_1 = b_2 = b \\ a_{11} = \gamma b - a_{22} \\ a_{12} = \gamma b - a_{21} \end{array} \right.$$

en hiervan gebruik makend vertaalt de symplecticiteit zich in het volgende stelsel:

$$\left\{ \begin{array}{l} \frac{b}{\gamma} a_{11} = b^2 - \frac{b}{\gamma} a_{11} \\ \frac{b}{\gamma} a_{21} = b^2 - \frac{b}{\gamma} a_{12} \\ \frac{b}{\gamma} a_{22} = b^2 - \frac{b}{\gamma} a_{22} \end{array} \right.$$

Deze twee stelsels worden vervolgens samengenomen en herschreven met behulp van twee andere parameters λ en θ . Deze weerspiegelen beter de symmetrie en maken de resultaten later duidelijker. De volgende coëfficiënten worden dan verkregen:

$$\left\{ \begin{array}{l} c_1 = \frac{1}{2} - \theta \\ c_2 = \frac{1}{2} + \theta \\ a_{11} = \frac{\gamma b}{2} \\ a_{12} = \frac{\gamma b}{2} + \lambda \\ a_{21} = \frac{\gamma b}{2} - \lambda \\ a_{22} = \frac{\gamma b}{2} \end{array} \right.$$

of in tableau-vorm:

$$\begin{array}{c|c|cc} \frac{1}{2} - \theta & \gamma & \frac{\gamma b}{2} & \frac{\gamma b}{2} + \lambda \\ \frac{1}{2} + \theta & \gamma & \frac{\gamma b}{2} - \lambda & \frac{\gamma b}{2} \\ \hline & & b & b \end{array} \quad (12)$$

Elke RK-methode integreert een constante functie exact (zie ook opmerking 1). Verder wordt op de interne trappen en de externe trap de functie $\exp(\pm\mu_1 x)$, met $\mu_1 = \omega_1 \vee i\omega_1, \omega_1 \in \mathbb{R}^+$, opgelegd, of met andere woorden:

$$S_{\text{int}} = \{\exp(\pm\mu_1 x)\}, \quad S_{\text{ext}} = \{1, \exp(\pm\mu_1 x)\}.$$

Met behulp van de functionalen \mathcal{L}_i en \mathcal{L} uit (5) toegepast op de functie $\exp(\pm\mu_1 x)$ worden de coëfficiënten:

$$b = \frac{\sinh(\frac{z_1}{2})}{\cosh(z_1\theta)z_1}, \quad \gamma = \frac{\cosh(2z_1\theta)}{\cosh(\frac{z_1}{2})\cosh(z_1\theta)}, \quad \lambda = -\frac{\sinh(z_1\theta)}{\cosh(z_1\theta)z_1}. \quad (13)$$

Hierbij is $z_1 = \mu_1 h$. Er is dus nog één vrijheidsgraad, θ , die kan vastgelegd worden door de functie $\exp(\pm\mu_2 x)$, met opnieuw $\mu_2 = \omega_2 \vee i\omega_2, \omega_2 \in \mathbb{R}^+$, op te leggen voor de externe trap:

$$S_{\text{int}} = \{\exp(\pm\mu_1 x)\}, \quad S_{\text{ext}} = \{1, \exp(\pm\mu_1 x), \exp(\pm\mu_2 x)\}.$$

Dit levert dezelfde voorwaarde op voor b in $z_2 = \mu_2 h$, waardoor θ kan bepaald worden uit:

$$\frac{\sinh(\frac{z_1}{2})}{\cosh(z_1\theta)z_1} = \frac{\sinh(\frac{z_2}{2})}{\cosh(z_2\theta)z_2}.$$

Met behulp van de eta-functies van Ixaru kan dit ook als volgt geschreven worden:

$$\frac{\eta_0(\frac{Z_1}{4})}{\eta_{-1}(Z_1\theta^2)} = \frac{\eta_0(\frac{Z_2}{4})}{\eta_{-1}(Z_2\theta^2)}. \quad (14)$$

Definiëren we nu

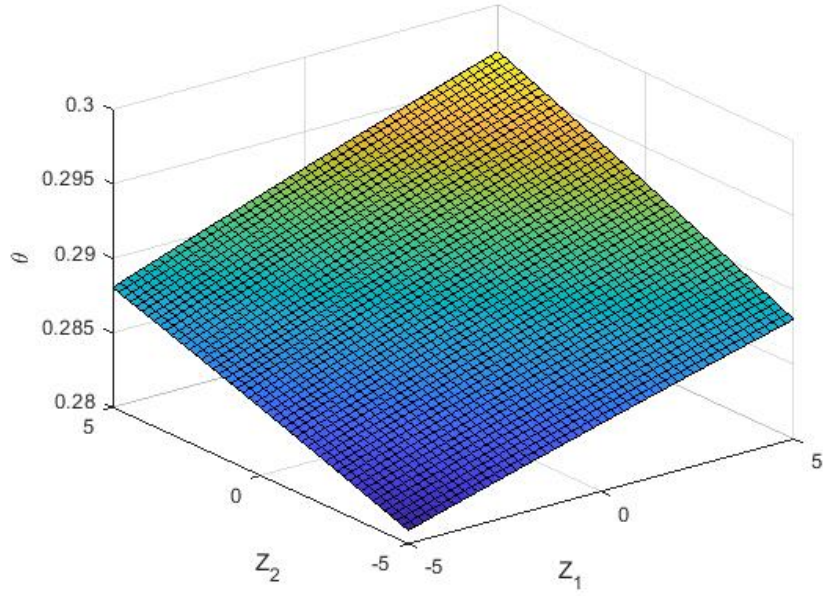
$$F(Z, \theta) = \frac{\eta_0(\frac{Z}{4})}{\eta_{-1}(Z\theta^2)},$$

dan kan de voorwaarde eenvoudig worden voorgesteld als $F(Z_1, \theta) - F(Z_2, \theta) = 0$. Er bestaat jammer genoeg geen oplossing voor deze vergelijking in expliciete vorm, dus de bepaling van θ zal via numerieke benaderingen moeten gebeuren. Een mogelijke manier hierbij is om te werken met de verhouding van beide frequenties:

$$\alpha = \frac{Z_2}{Z_1} = \frac{\mu_2^2}{\mu_1^2}.$$

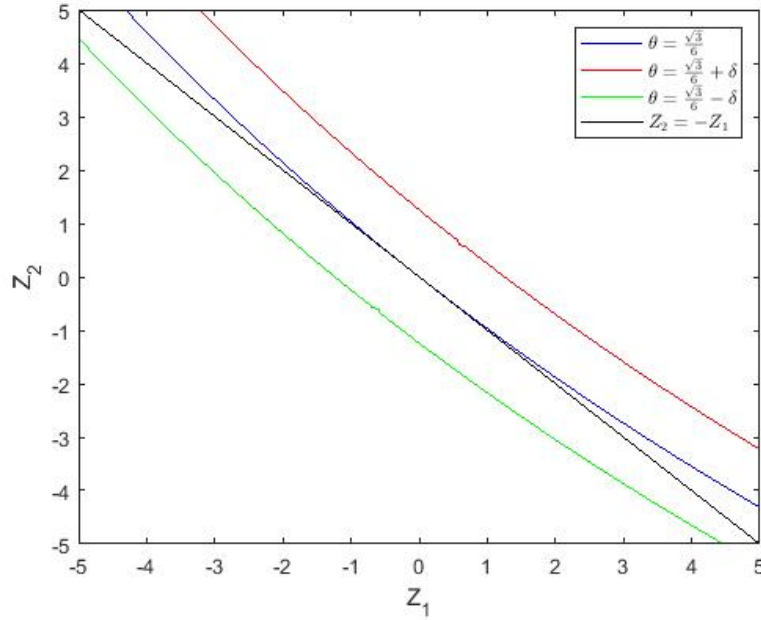
Bij positieve waarden van α zijn de frequenties allebei reëel of complex, bij negatieve waarden komt zowel een reële als een complexe frequentie voor. Voor verschillende waarden van α kan de vergelijking $G(Z, \theta) = F(Z, \theta) - F(\alpha Z, \theta) = 0$ numeriek worden opgelost. Het resultaat hiervan voor $Z_i \in [-5, 5]$ wordt geplot in figuur 1. Merk op dat de vergelijking voor θ symmetrisch is in beide frequenties en dit dus ook weerspiegeld wordt in deze grafiek.

Notatie 14. *In de numerieke resultaten zal voor deze methode de naam $RKs2$ gebruikt worden.*



Figuur 1: θ in functie van Z_1 en Z_2 .

Vaak wordt ook de vraag gesteld wat er gebeurt wanneer de integratiepunten c_i op voorhand vast worden genomen, meestal op de gaussische knooppunten, wat correspondeert met $\theta = \frac{\sqrt{3}}{6}$ in het geval van de tweetrapsmethode. Gezien hierdoor een vrijheidsgraad verloren is, kan er niet meer gefit worden op een extra frequentie μ_2 , maar impliciet legt de keuze van een constante θ toch een tweede z_2 vast die door de externe trap exact wordt geïntegreerd: de snijlijn van het impliciete oppervlak uit figuur 1 en het vlak $\theta = \frac{\sqrt{3}}{6}$. Voor kleine waarden van z_1 leunt z_2 nauw aan tegen iz_1 , waardoor het lijkt alsof de externe trappen zowel een exponentiële als een goniometrische component exact integreren. Dit alles wordt geïllustreerd in figuur 2: de blauwe kromme is de vermelde doorsnede en ligt inderdaad dicht bij het geval $Z_2 = -Z_1$, equivalent met $z_2 = \pm iz_1$. Een uitgebreide behandeling van deze versie van de methode kan teruggevonden worden in [15].

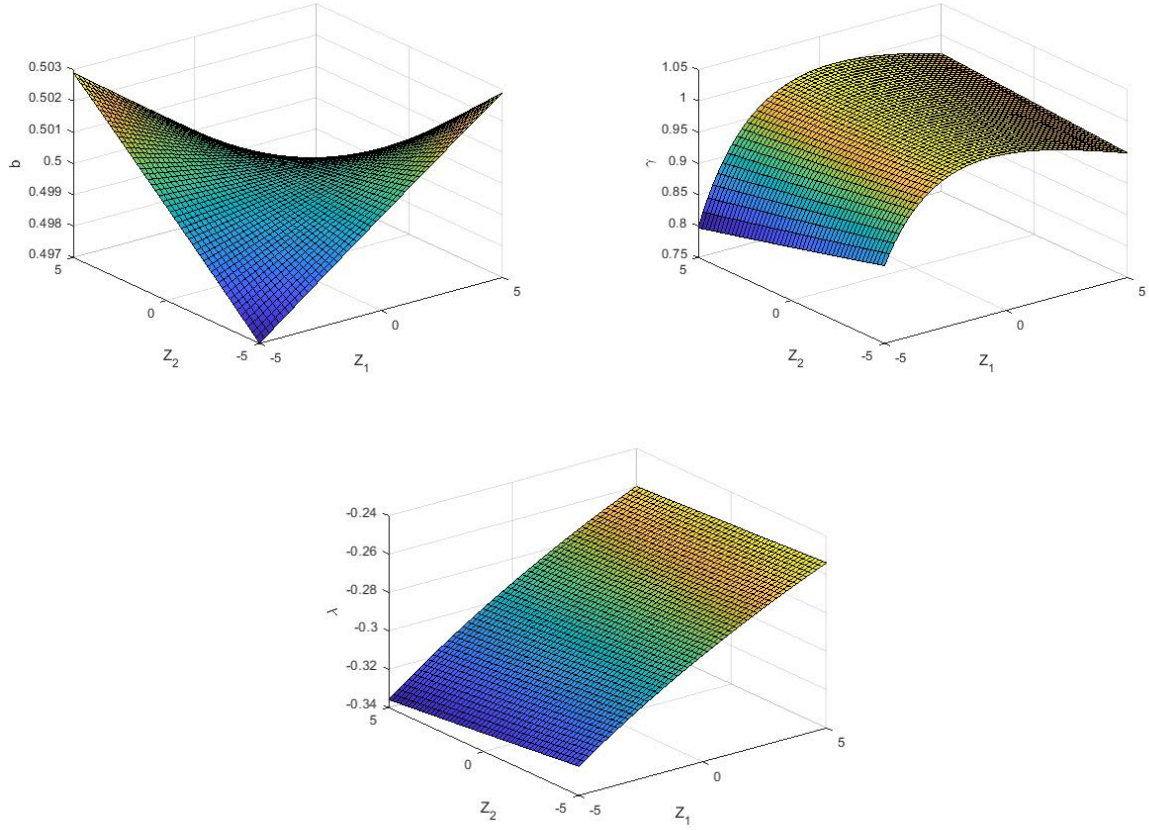


Figuur 2: De ligging van $\theta = \frac{\sqrt{3}}{6}$ en verwante waarden, met $\delta = 0.001$.

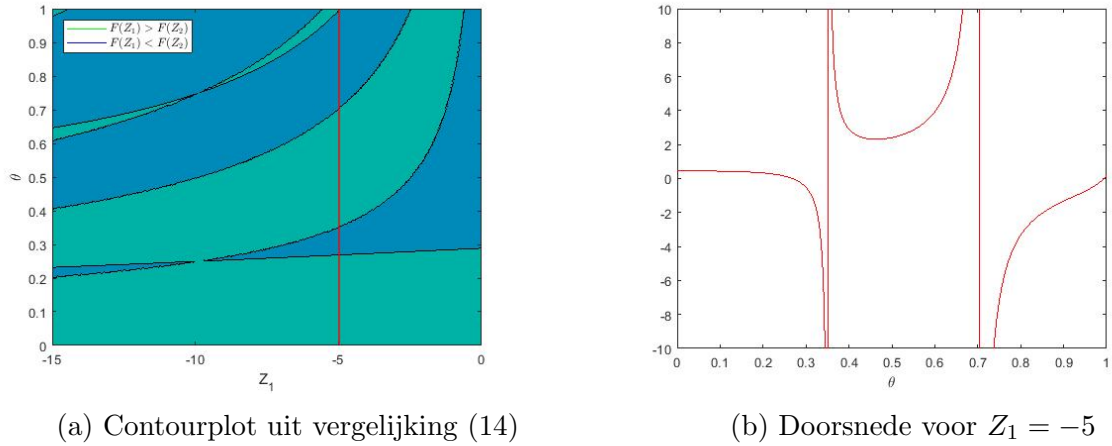
3.2 Coëfficiënten

Een eerste observatie omtrent de coëfficiënten volgt al uit figuur 1: θ lijkt als functie van z_1 en z_2 niet heel variabel. Omdat de z_i vaak schattingen zijn met een zekere foutmarge, is dit een zeer handige eigenschap. Naarmate de verhouding α in absolute waarde toeneemt, neemt variatie wel toe. De andere coëfficiënten, geplotted in hetzelfde interval als θ in figuur 3, zijn iets meer variabel, voornamelijk γ voor sterk negatieve Z_1 , maar verder vertonen ze geen problematisch gedrag.

Er ontstaat echter een ander probleem bij negatieve Z_i : in principe is θ bepaald over geheel \mathbb{R}^2 , maar in het geval dat een Z_i negatief wordt en de corresponderende z_i dus imaginair, komen er in de bepaling van de parameters cosinussen met nulpunten. Dit wordt geïllustreerd in figuur 4a, waar de nulpunten worden geplotted van vergelijking (14) voor $z_2 = 2z_1$ met behulp van een contourplot. Slechts drie van de vijf krommen op de figuur stellen wel degelijk nulpunten voor, de twee andere lijnen zijn fouten ten gevolge van de polen van de vergelijking (ontstaan door de nulpunten van de cosinussen in de noemer). Figuur 4b, die een doorsnede weergeeft voor het geval $Z_1 = -5$, verduidelijkt dit: de rode kromme heeft slechts twee nulpunten in het interval $[0, 1]$ en twee verticale asymptoten. Wanneer nu een dergelijke poollijn uit het contourplot snijdt met een werkelijke contour, dan ontstaan er soms sprongen naar een andere contour bij de numerieke bepaling van θ . In theorie kunnen verschillende van deze discontinuïteiten relatief eenvoudig weggewerkt worden, met behulp van l'Hopital, maar een aparte numerieke behandeling bij al deze gevallen is allerm minst wenselijk. Bovendien treden dergelijke discontinuïteiten ook op bij de coëfficiënten en naarmate de verhouding $|\alpha|$ groter wordt, treden ze zelfs eerder op. Bijgevolg is de methode praktisch niet meer bruikbaar voor grote imaginaire z_i . Dit alles maakt dat het toepassen van de methode best beperkt blijft tot relatief kleine frequenties.



Figuur 3: Het verloop van de verschillende coëfficiënten in functie van de gefitte parameters.



Figuur 4: Illustratie van de nulpunten en polen bij de bepaling van θ .

Zoals bij alle exponentieel gefitte methoden kunnen kleine waarden van z_i zorgen voor fouten bij het berekenen van de coëfficiënten. De frequenties komen immers in teller en noemer van alle uitdrukkingen voor, waardoor kleine getallen met afrondingsfouten door

elkaar gedeeld worden. Daarom is het beter om vanaf een zekere grenswaarde voor $|z_i|$, vaak 0.1 gekozen, de reeksontwikkelingen van de coëfficiënten te gebruiken. Voor θ geeft dit het volgende resultaat:

$$\begin{aligned}\theta = & \frac{\sqrt{3}}{6} + \frac{\sqrt{3}}{2160}(z_1^2 + z_2^2) - \frac{\sqrt{3}}{10886400}(27z_1^4 - 106z_1^2z_2^2 + 27z_2^4) \\ & + \frac{\sqrt{3}}{435456000}(3z_1^4 - 34z_1^2z_2^2 + 3z_2^4)(z_1^2 + z_2^2) + \dots\end{aligned}\quad (15)$$

Met behulp hiervan kunnen ook de andere drie parameters benaderd worden met een reeksontwikkeling:

$$\begin{aligned}b = & \frac{1}{2} + \frac{z_1^2(z_1^2 - 252)z_2^2}{2177280} + \frac{(-244500z_1^6 + 21168000z_1^4 + 435456000z_1^2)z_2^4}{948109639680000} \\ & + \frac{(-3943z_1^6 - 91320z_1^4)z_2^6}{948109639680000} + \dots\end{aligned}\quad (16)$$

$$\begin{aligned}\lambda = & -\frac{\sqrt{3}}{6} + \frac{\sqrt{3}z_1^2}{240} - \frac{137\sqrt{3}z_1^4}{1209600} + \frac{143\sqrt{3}z_1^6}{48384000} + \left(-\frac{\sqrt{3}}{2160} + \frac{157\sqrt{3}z_1^2}{5443200} - \frac{1367\sqrt{3}z_1^4}{1306368000}\right)z_2^2 \\ & + \left(\frac{\sqrt{3}}{403200} - \frac{37\sqrt{3}z_1^2}{1306368000}\right)z_2^4 - \frac{\sqrt{3}z_2^6}{145152000} + \dots\end{aligned}\quad (17)$$

$$\gamma = 1 - \frac{z_1^4}{360} + \frac{11z_1^6}{30240} + \left(\frac{z_1^2}{1440} - \frac{29z_1^4}{362880}\right)z_2^2 - \frac{z_1^2z_2^4}{362880} + \dots\quad (18)$$

Wegens het multiparametrisch karakter van de methode is het nu echter mogelijk dat slechts één van beide frequenties onder de grenswaarde ligt. Hier kunnen reeksontwikkelingen in enkel de probleemfrequentie gebruikt worden. Indien in het bijzonder enkel z_2 zeer klein is, dient enkel een benadering gebruikt te worden voor θ . De overige coëfficiënten, gedefinieerd in (13), zijn immers enkel onrechtstreeks afhankelijk van z_2 via θ .

3.3 Foutterm

De orde van methode (12) volgt uit lemma 9: de laagste-ordetermen in h treden op vanaf orde vijf, zowel bij de sommatie over de gewone bomen als over de spookbomen. Hieronder worden alle bomen tot en met orde vijf en de spookbomen overlopen en wordt de bijdrage aan de foutterm berekend.

3.3.1 Bomen van orde één

Er is slechts één boom van orde één, namelijk τ . De factor $\frac{h}{1!}$ draagt al één bij om tot een vijfde-ordeterm te komen, bijgevolg dient op zoek gegaan te worden naar de vierde-ordeterm in h van de reeksontwikkeling van $\psi(\tau)$. Dit geeft als resultaat

$$\frac{h^5}{4320}\mu_1^2\mu_2^2f,$$

of in termen van gewone afgeleiden

$$\frac{h^5}{4320}\mu_1^2\mu_2^2y^{(1)}.$$

3.3.2 Bomen van orde twee

De orde van h voorop in de sommatie is nu twee, waardoor een mogelijke bijdrage aan de vijfde-ordeterm nu moet komen van de derde-ordeterm van de reeksontwikkeling van $\psi(t)$, met t een boom van orde twee. Deze reeksontwikkeling is echter zoals eerder afgeleid even, dus deze bomen hebben geen effect op de lte.

3.3.3 Bomen van orde drie

Uit een analoge redenering als bij de bomen van orde een volgt dat op zoek dient gegaan te worden naar de tweede-ordetermen van de reeksontwikkelingen van $\psi(t)$, $\rho(t) = 3$. De bomen en hun corresponderende term zijn de volgende:

- $[\tau^2] \Rightarrow \frac{h^5}{4320} (9\mu_1^2 - \mu_2^2) \{f^2\}$
- $[2\tau]_2 \Rightarrow \frac{h^5}{2160} (\mu_2^2 - 9\mu_1^2) \{2f\}_2$

Groeperen in termen van gewone afgeleiden geeft:

$$\frac{h^5}{4320} (9\mu_1^2 - \mu_2^2) (y^{(3)} - 3Jy^{(2)}).$$

3.3.4 Bomen van orde vier

Net als bij orde twee is de relevante term, nu die van eerste orde, gelijk aan nul.

3.3.5 Bomen van orde vijf

Nu de factor voorop al h^5 is, dienen enkel nog de constante termen in de ontwikkelingen van $\psi(t)$ berekend te worden, met als resultaat:

- $[\tau^4] \Rightarrow \frac{h^5}{4320} \{f^4\}$
- $[[\tau]^2] \Rightarrow \frac{h^5}{1440} \{\{f\}^2\}$
- $[\tau[\tau^2]] \Rightarrow -\frac{h^5}{720} \{f\{f^2\}\}$
- $[\tau[2\tau]_2] \Rightarrow -\frac{h^5}{720} \{f\{2f\}_2\}$
- $[\tau^2[\tau]] \Rightarrow \frac{h^5}{720} \{f^2\{f\}\}$
- $[2\tau^3]_2 \Rightarrow -\frac{h^5}{1080} \{2f^2\}$
- $[2\tau[\tau]]_2 \Rightarrow -\frac{h^5}{360} \{2f\{f\}\}_2$
- $[3\tau^2]_3 \Rightarrow \frac{h^5}{720} \{3f^2\}_3$
- $[4\tau]_4 \Rightarrow \frac{h^5}{720} \{4f\}_4$

Sommeren en herschrijven geeft

$$\frac{h^5}{4320} (10J^2y^{(3)} - 5Jy^{(4)} + y^{(5)} - 10\mathbf{f}(y^{(1)}, y^{(3)})),$$

waarbij \mathbf{f} de tweede Fréchetafgeleide van f voorstelt.

3.3.6 Spookbomen

De relevante spookbomen t zijn diegene waarvoor $\rho(t) + \rho_g(t) \leq 4$, of dus $[\tau_g]$, $[_2\tau_g]_2$ en $[\tau\tau_g]$. Voor de twee bomen van orde drie, $[_2\tau_g]_2$ en $[\tau\tau_g]$, geldt echter dat $\rho'(t) = 2$ en dus dat de orde van h voorop in de term twee is. Net zoals bij de gewone bomen van even orde betekent dit dat de relevante ordetermen nul zijn en dus blijft enkel $[\tau_g]$ over. De spookterm die daaruit volgt is

$$\frac{h^5}{1440} (4\mu_1^4 - \mu_1^2\mu_2^2) f = \frac{h^5}{1440} (4\mu_1^4 - \mu_1^2\mu_2^2) y^{(1)}.$$

3.3.7 Interpretatie

Indien we al deze termen bij elkaar nemen, dan wordt de volgende uitdrukking voor de foutterm verkregen:

$$\begin{aligned} y(x_n) - y_n = & \frac{h^5}{4320} \left((12\mu_1^4 - 2\mu_1^2\mu_2^2)y^{(1)} + (9\mu_1^2 - \mu_2^2)(y^{(3)} - 3Jy^{(2)}) \right. \\ & \left. + 10J^2y^{(3)} - 5Jy^{(4)} + y^{(5)} - 10\mathbf{f}(y^{(1)}, y^{(3)}) \right) + \mathcal{O}(h^6), \end{aligned} \quad (19)$$

Met behulp hiervan zou nu op zoek gegaan kunnen worden naar de optimale parameters μ_i , gegeven een functie y : zoek goeie benaderingen voor $y^{(2)}, y^{(3)}, y^{(4)}$ en $y^{(5)}$ om vervolgens de leidende foutterm te minimaliseren over alle mogelijke parameterwaarden. In de realiteit zijn dergelijke benaderingen moeilijk te vinden, omdat de uitdrukking voor y niet gekend is. Toch kunnen in (19) al twee dingen herkend worden:

- Indien $\mu_1 = \mu_2 = 0$ worden gekozen, komt de foutterm overeen met die van de klassieke gaussische tweetrapsmethode, zoals gevonden in [17].
- Neem de vergelijking $y' = \omega y$, met als oplossing $y = \exp(\omega x)$, dan vereenvoudigt (19) tot

$$y(x_n) - y_n = \frac{h^5}{2160} (6\mu_1^4 - \mu_1^2\mu_2^2 - 9\mu_1^2\omega^2 + \mu_2^2\omega^2 + 3\omega^4)y^{(1)} + \mathcal{O}(h^6).$$

De middelste factor van de lte factoriseert verder tot $(\mu_1^2 - \omega^2)(\mu_1^2 - \frac{\mu_2^2 + 3\omega^2}{6})$, dus $\mu_1 = \pm\omega$ maakt de leidende foutterm nul. Ook de andere fouttermen zullen nul worden, consistent met het feit dat de interne en externe trappen de frequentie μ_1 exact integreren.

Bij de numerieke voorbeelden in hoofdstuk 6 worden verdere pogingen ondernomen om de foutterm te interpreteren.

3.4 Stabiliteit

De stabiliteitsfunctie wordt met behulp van (10):

$$R(z) = \frac{1 + \gamma bz + \lambda^2 z^2}{1 - \gamma bz + \lambda^2 z^2}. \quad (20)$$

De grens van het gebied van absolute stabiliteit \mathcal{R}_A wordt als volgt bepaald (met $z = x + iy$):

$$\begin{aligned}
& |R(z)| = 1 \\
& \Leftrightarrow |1 + \gamma bz + \lambda^2 z^2| = |1 - \gamma bz + \lambda^2 z^2| \\
& \Leftrightarrow |1 + \gamma bz + \lambda^2 z^2|^2 = |1 - \gamma bz + \lambda^2 z^2|^2 \\
& \Leftrightarrow |(1 + \gamma bx + \lambda^2(x^2 - y^2)) + iy(\gamma b + 2x\lambda^2)|^2 = |(1 - \gamma bx + \lambda^2(x^2 - y^2)) + iy(2x\lambda^2 - \gamma b)|^2 \\
& \Leftrightarrow (1 + \gamma bx + \lambda^2(x^2 - y^2))^2 + y^2(\gamma b + 2x\lambda^2)^2 = (1 - \gamma bx + \lambda^2(x^2 - y^2))^2 + y^2(2x\lambda^2 - \gamma b)^2 \\
& \Leftrightarrow 2\gamma bx + 2\gamma bx\lambda^2(x^2 - y^2) + 4\gamma bx\lambda^2 y^2 = -2\gamma bx - 2\gamma bx\lambda^2(x^2 - y^2) - 4\gamma bx\lambda^2 y^2 \\
& \Leftrightarrow x + x\lambda^2(x^2 - y^2) + 2x\lambda^2 y^2 = 0 \\
& \Leftrightarrow x(1 + \lambda^2(x^2 + y^2)) = 0
\end{aligned}$$

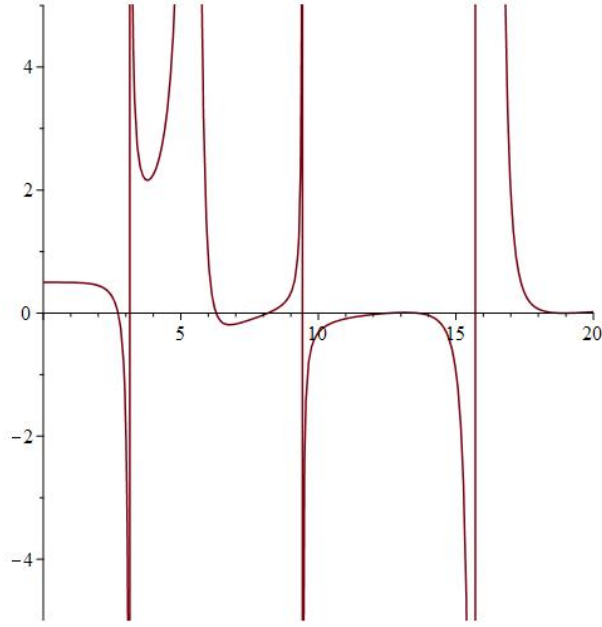
De tweede factor in de finale gelijkheid is strikt positief, waardoor de oplossing $x = 0$ wordt, of dus de imaginaire as. Bij deze berekening is enkel gesteund op het feit dat het product γb niet nul is, wat wel het geval kan zijn voor imaginaire z_i : er komt dan een sinus in de teller van b en een cosinus in de teller van γ en beide hebben nulpunten, gegeven door:

$$z_1 = \frac{(2j+1)\pi i}{4\theta}, \quad j \in \mathbb{Z} \quad \vee \quad z_1 = 2j\pi i, \quad j \in \mathbb{Z}^0.$$

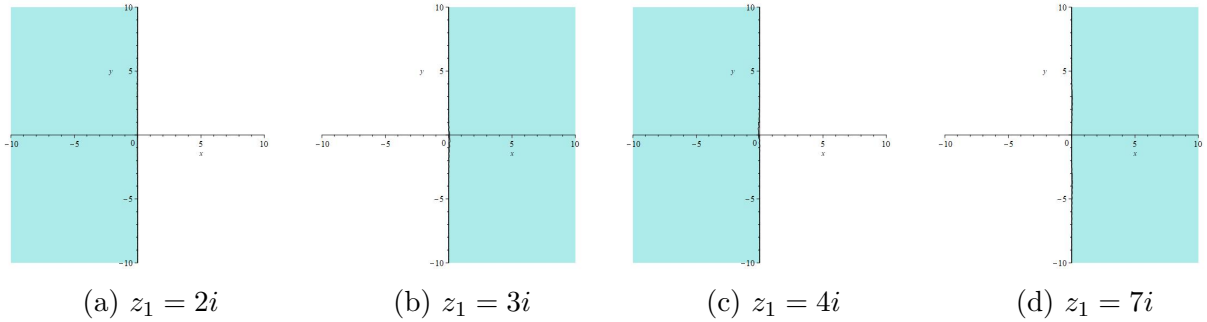
In deze nulpunten vereenvoudigt de stabiliteitsfunctie tot $R(z) = 1$, onafhankelijk van z , en het stabiliteitsgebied wordt \emptyset . Dit resulteert dus zelfs niet meer in een methode die een veralgemening is van een klassieke eerste-orde-methode (waarvoor $R(z)$ dus $\exp(z)$ minstens tot op eerste orde benadert). De tweetrapsmethode wordt dus nutteloos bij dergelijke keuzes voor z_i en bovendien ontstaan er nog problematische gevallen. Om te kijken waar ten opzichte van de imaginaire as het stabiliteitsgebied nu precies ligt, kan gekeken worden naar $|R(-1)|$, of dus:

$$\left| \frac{\gamma b - \lambda^2 - 1}{\gamma b + \lambda^2 + 1} \right|.$$

Deze waarde is kleiner dan één (en -1 ligt dus in \mathcal{R}_A) als en slechts als γb positief is. Hieruit volgt dus dat het stabiliteitsgebied \mathbb{C}^- is als en slechts als γb positief is. In figuur 5 wordt dit product weergegeven voor $z_1 \in [0, 20i]$, waarop duidelijk negatieve zones te zien zijn. Hierdoor wisselt bij elke overgang van positief naar negatief het stabiliteitsgebied van \mathbb{C}^- naar \mathbb{C}^+ en omgekeerd, zoals geïllustreerd in figuur 6.



Figuur 5: Het verloop van γb voor $z_1 \in [0, 20i]$ en $\theta = \frac{\sqrt{3}}{6}$.



Figuur 6: Stabiliteitsgebieden (turquoise) voor verschillende waarden van z_1 .

Voor reële frequenties wordt het gebied van absolute stabiliteit dus steeds gegeven door \mathbb{C}^- en is de methode onvoorwaardelijk exact-A-stabiël. Merk op dat hier de stabiliteitseigenschap van de klassieke gaussische methode, die eveneens exact-A-stabiël is, wordt overgeërfd. Voor imaginaire frequenties hangt de stabiliteit sterk af van de z_i en schommelt de methode tussen exacte A-stabiliteit, exacte A-instabiliteit en helemaal geen stabiliteit. Omdat deze schommelingen niet voorkomen bij kleine waarden van $|z_i|$, is het dus ook om stabiliteitsredenen verstandig om het gebruik van de tweetrapsmethode te beperken tot relatief kleine frequenties.

3.5 Speciale gevallen

Voor enkele specifieke keuzes van z_2 worden reeds gekende methodes verkregen. Wanneer bijvoorbeeld beide frequenties gelijk worden gekozen, resulterend in $z_2 = z_1$, dan wordt de

voorwaarde ter bepaling van θ gegeven door $\frac{\partial F(Z, \theta)}{\partial Z} \Big|_{Z=Z_1} = 0$, wat vereenvoudigt tot:

$$\theta = \frac{\cosh(z_1 \theta)}{z_1 \sinh(z_1 \theta)} \left(\frac{z_1 \cosh(\frac{z_1}{2})}{2 \sinh(\frac{z_1}{2})} - 1 \right). \quad (21)$$

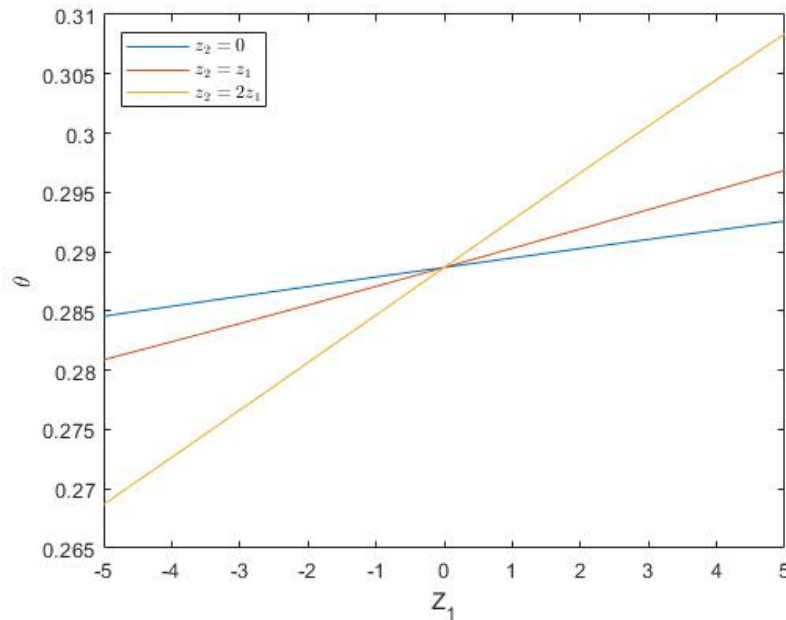
Deze specifieke versie werd al eerder onderzocht door Van den Berghe et al. in [16]. Een tweede optie die hierin voorkomt is die waarbij de tweede frequentie wordt ingeruild voor een hogere polynomiale orde, corresponderend met $z_2 = 0$. Aangezien $F(0, \theta) = 1$, wordt de vergelijking ter bepaling van θ nu $F(Z_1, \theta) = 1$, waarvoor een expliciete oplossing bestaat:

$$\theta = \frac{1}{z_1} \arccos \left(\frac{2 \sinh(\frac{z_1}{2})}{z_1} \right). \quad (22)$$

Het geval waarbij de tweede frequentie het dubbele is van de eerste, of dus $z_2 = 2z_1$, is nog niet specifiek onderzocht, maar geeft ook een exacte oplossing voor θ :

$$\theta = \frac{1}{z_1} \arccos \left(\frac{\cosh(\frac{z_1}{2}) + \sqrt{8 + \cosh(\frac{z_1}{2})^2}}{4} \right). \quad (23)$$

Het verloop van θ wordt voor elk van deze speciale gevallen weergegeven in figuur 7, die nogmaals illustreert dat een grotere $|\alpha|$ voor een variabelere θ zorgt, en dus een nauwkeurigere schatting van de z_i vereist. Het spreekt voor zich dat alle resultaten die tot nu toe zijn bekomen eveneens gelden voor deze specifieke gevallen door hun respectievelijke keuzes van z_2 in te voegen.



Figuur 7: Het verloop van θ voor enkele speciale gevallen van z_2 .

4 Drietrapsmethode

4.1 De methode

Een drietraps gemodificeerde RK-methode is van de vorm:

$$\begin{array}{c|ccc} c_1 & \gamma_1 & a_{11} & a_{12} & a_{13} \\ c_2 & \gamma_2 & a_{21} & a_{22} & a_{23} \\ c_3 & \gamma_3 & a_{31} & a_{32} & a_{33} \\ \hline & & b_1 & b_2 & b_3 \end{array}$$

Als hierop symmetrievorwaarden worden opgelegd voldoen de coëfficiënten aan het volgende stelsel:

$$\left\{ \begin{array}{l} c_1 = 1 - c_3 \\ c_2 = \frac{1}{2} \\ \gamma_1 = \gamma_3 \\ b_1 = b_3 \\ a_{11} = \gamma_1 b_1 - a_{33} \\ a_{12} = \gamma_1 b_2 - a_{32} \\ a_{13} = \gamma_1 b_1 - a_{31} \\ a_{21} = \gamma_2 b_1 - a_{23} \\ a_{22} = \frac{\gamma_2 b_2}{2} \end{array} \right.$$

en de bijkomende voorwaarden die uit symplecticiteit volgen zijn:

$$\left\{ \begin{array}{l} a_{11} = \frac{b_1 \gamma_1}{2} \\ \frac{b_2}{\gamma_2} a_{21} = b_1 b_2 - \frac{b_1}{\gamma_1} a_{12} \\ \frac{b_1}{\gamma_1} a_{32} = b_1 b_2 - \frac{b_2}{\gamma_2} a_{23} \\ a_{33} = \frac{b_1 \gamma_1}{2} \end{array} \right.$$

Indien we opnieuw enkele andere parameters, θ , α_2 en α_3 , invoeren die de resultaten later duidelijker gaan maken en de voorwaarden verder vereenvoudigen, bekomen we de volgende methode in tableau-vorm:

$$\begin{array}{c|ccc} \frac{1}{2} - \theta & \gamma_1 & \frac{\gamma_1 b_1}{2} & \frac{\gamma_1 b_2}{2} - \alpha_2 & \frac{\gamma_1 b_1}{2} - \alpha_3 \\ \frac{1}{2} & \gamma_2 & \frac{\gamma_2 b_1}{2} + \frac{b_1 \alpha_2 \gamma_2}{\gamma_1 b_2} & \frac{\gamma_2 b_2}{2} & \frac{\gamma_2 b_1}{2} - \frac{b_1 \alpha_2 \gamma_2}{\gamma_1 b_2} \\ \frac{1}{2} + \theta & \gamma_1 & \frac{\gamma_1 b_1}{2} + \alpha_3 & \frac{\gamma_1 b_2}{2} + \alpha_2 & \frac{\gamma_1 b_1}{2} \\ \hline & & b_1 & b_2 & b_1 \end{array} \quad (24)$$

Om tot concrete waarden te komen, kan er opnieuw gestart worden met het opleggen van de functie $\exp(\pm \mu_1 x)$, $\mu_1 \in \mathbb{C}$, voor alle trappen (bovenop de constante functie die opnieuw standaard is), of dus:

$$S_{\text{int}} = \{\exp(\pm \mu_1 x)\}, \quad S_{\text{ext}} = \{1, \exp(\pm \mu_1 x)\}.$$

Het oplossen van deze voorwaarden levert nu echter nog geen eenvoudige uitdrukkingen op in één resterende parameter θ , zoals in het tweetrapsgeval, en is ook nog afhankelijk van de parameters γ_1 en γ_2 . Om die laatste twee weg te werken zijn dus bijkomende voorwaarden nodig, en de meest voor de hand liggende is eisen dat de constante functie ook exact wordt geïntegreerd door de interne trappen: $S_{\text{ext}} = \{1, \exp(\pm\mu_1 x)\}$. Hiermee worden de coëfficiënten:

$$\begin{aligned} \gamma_1 &= 1, & \gamma_2 &= 1, & b_1 &= \frac{\frac{\sinh(z_1)}{z_1} - \frac{\sinh(\frac{z_1}{2})}{\frac{z_1}{2}}}{2(\cosh(2z_1\theta) - \cosh(z_1\theta))}, \\ b_2 &= \frac{\sinh(z_1)}{z_1 \cosh(\frac{z_1}{2})} - 2b_1 \cosh(z_1\theta), & \alpha_2 &= \frac{2 \sinh(\frac{z_1}{2}) \cosh(2z_1\theta) - \sinh(z_1) \cosh(z_1\theta)}{2z_1 \sinh(z_1\theta) \sinh(\frac{z_1}{2})}, \\ \alpha_3 &= \frac{\sinh(z_1) - 2 \sinh(\frac{z_1}{2}) \cosh(z_1\theta)}{2z_1 \sinh(z_1\theta) \sinh(\frac{z_1}{2})}, \end{aligned}$$

waarbij $z_1 = \mu_1 h$. De laatste vrijheidsgraad θ kan opnieuw worden bepaald door het opleggen van een extra frequentie μ_2 voor de externe trap:

$$S_{\text{int}} = \{1, \exp(\pm\mu_1 x)\}, \quad S_{\text{ext}} = \{1, \exp(\pm\mu_1 x), \exp(\pm\mu_2 x)\}.$$

Dit resulteert in een extra vergelijking in de b_i 's, waaruit een nieuwe uitdrukking voor b_2 in functie van b_1 wordt gehaald. Indien die wordt gelijkgesteld aan de eerder bekomen uitdrukking voor b_2 , kan daaruit b_1 worden opgelost:

$$b_1 = \frac{\frac{\sinh(\frac{z_2}{2})}{\frac{z_2}{2}} - \frac{\sinh(\frac{z_1}{2})}{\frac{z_1}{2}}}{2(\cosh(z_2\theta) - \cosh(z_1\theta))} = \frac{\eta_0\left(\frac{Z_2}{4}\right) - \eta_0\left(\frac{Z_1}{4}\right)}{2(\eta_{-1}(Z_2\theta^2) - \eta_{-1}(Z_1\theta^2))}.$$

Omdat deze uitdrukking van dezelfde vorm is als de eerste voorwaarde op b_1 , voeren we volgende functie in:

$$G(a, b) = \frac{\frac{\sinh(\frac{a}{2})}{\frac{a}{2}} - \frac{\sinh(\frac{b}{2})}{\frac{b}{2}}}{\cosh(a\theta) - \cosh(b\theta)} = \frac{\eta_0\left(\frac{a^2}{4}\right) - \eta_0\left(\frac{b^2}{4}\right)}{2(\eta_{-1}(a^2\theta^2) - \eta_{-1}(b^2\theta^2))}.$$

De gelijkheid van b_1 voor beide frequenties komt nu neer op de volgende vergelijking:

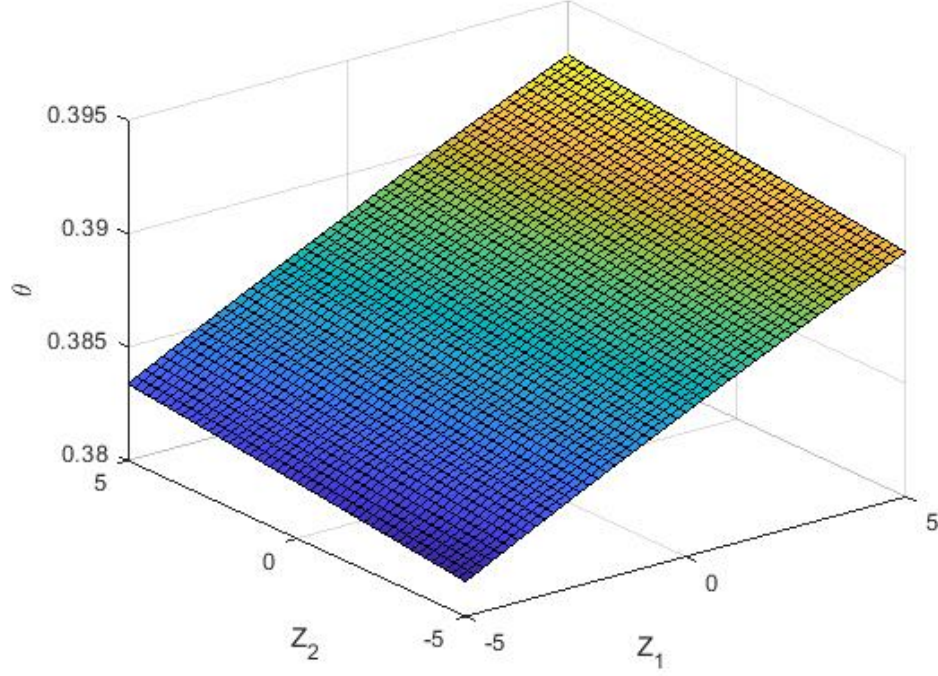
$$G(z_1, z_2) = G(z_1, 2z_1). \quad (25)$$

Indien $z_2 = \pm 2z_1$ is hieraan automatisch voldaan, dus om θ te bepalen (en bijgevolg de methode volledig vast te leggen) dient te gelden dat $z_2 \neq \pm 2z_1$. Dit houdt dus ook in dat de werkelijke fitting space van de volgende vorm is:

$$S_{\text{int}} = \{1, \exp(\pm\mu_1 x)\}, \quad S_{\text{ext}} = \{1, \exp(\pm\mu_1 x), \exp(\pm 2\mu_1 x), \exp(\pm\mu_2 x)\}.$$

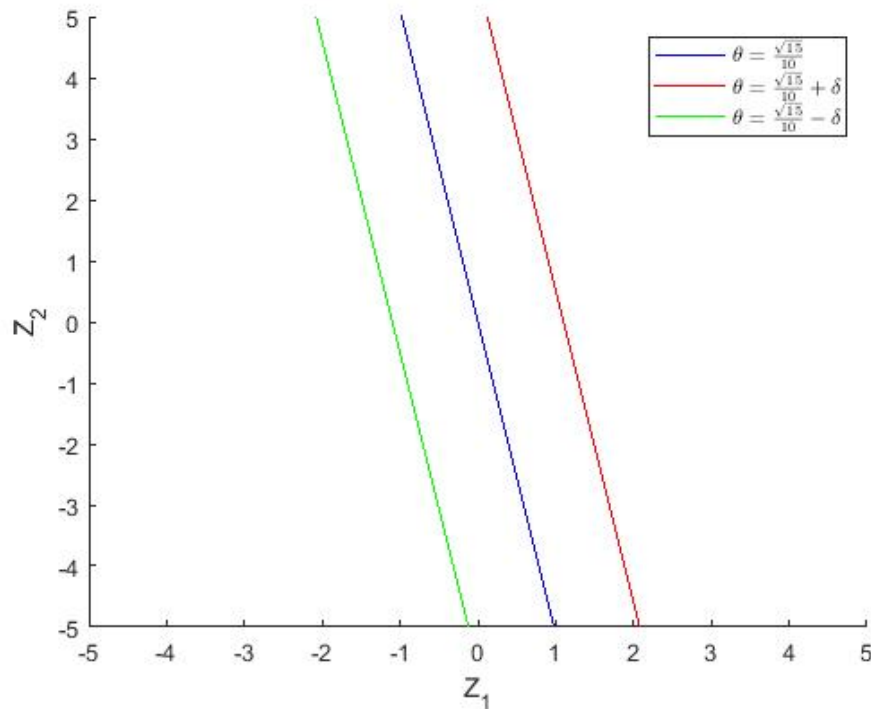
Er bestaat geen expliciete oplossing voor (25) in het algemene geval, maar de oplossingen kunnen eenvoudig numeriek worden teruggevonden. Het verloop van θ kan voor $Z_i \in [-5, 5]$ teruggevonden worden in figuur 8.

Notatie 15. *In de numerieke resultaten zal voor deze methode de naam $RKs3$ gebruikt worden.*



Figuur 8: θ in functie van Z_1 en Z_2

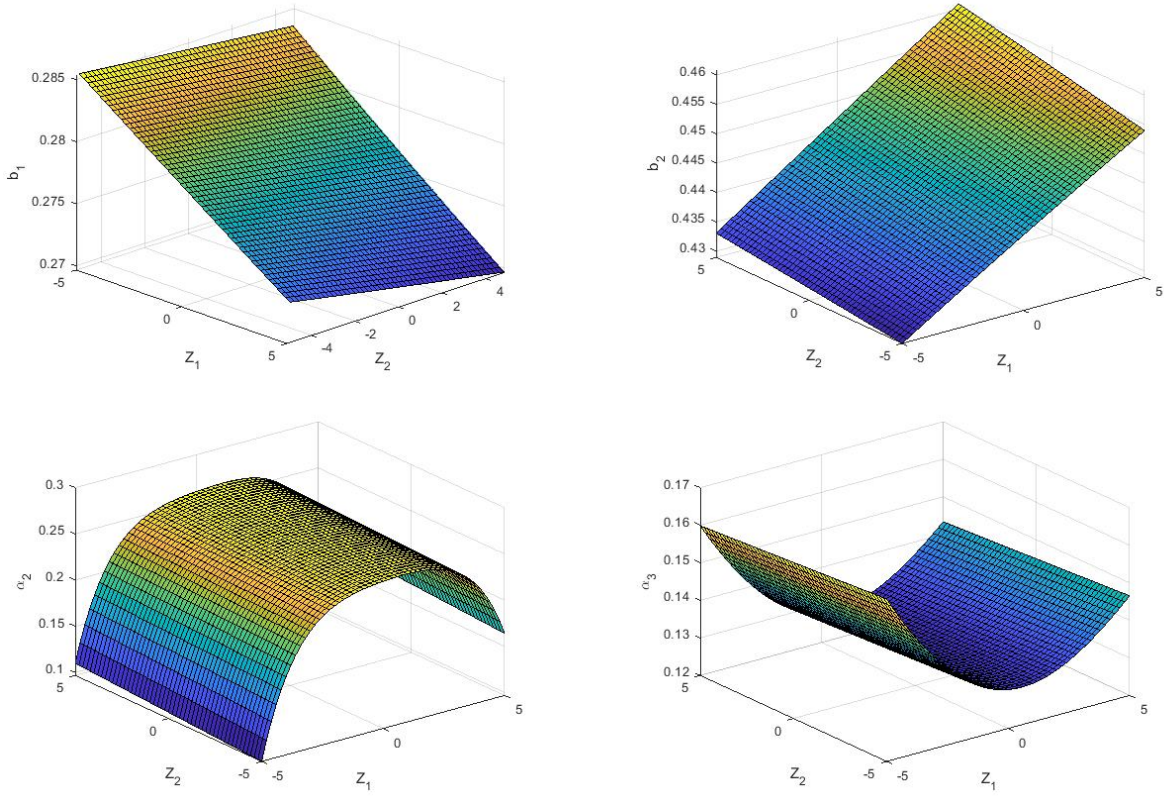
Opnieuw kan de vraag gesteld worden wat er gebeurt indien met de gaussische knooppunten gewerkt wordt. In het drietrapsgeval correspondeert dit met $\theta = \frac{\sqrt{15}}{10}$. Ook met deze keuze wordt impliciet een tweede frequentie vastgelegd, die kan gevonden worden door te kijken waar deze constante waarde van θ ligt op het oppervlak in figuur 8. In figuur 9 wordt de impliciete Z_2 weergegeven voor het gaussische geval en een koppel licht afwijkende waarden. Een verdere uitwerking kan in [2] gevonden worden.



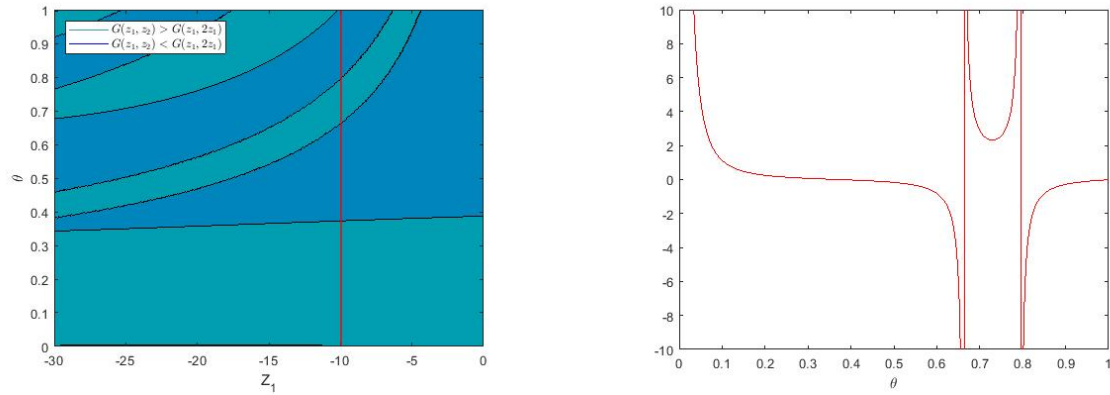
Figuur 9: De ligging van $\theta = \frac{\sqrt{15}}{10}$ en verwante waarden, met $\delta = 0.001$.

4.2 Coëfficiënten

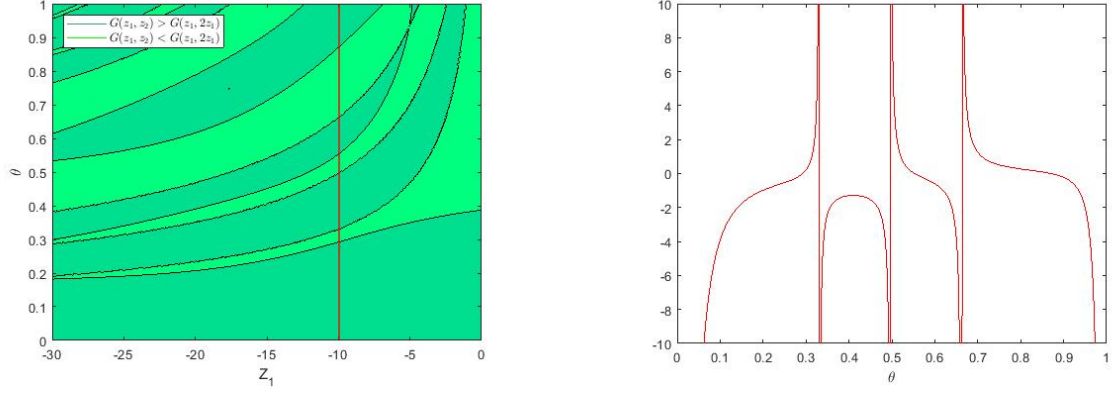
Uit figuur 8 kan in de eerste plaats dezelfde conclusie getrokken worden als in het tweetraps-geval: θ is weinig variabel. De symmetrie in de Z_i geldt niet langer, wat ook al uit de vorm van vergelijking (25) af te leiden was, en de grotere afhankelijkheid van z_1 in die vergelijking trekt zich eveneens door in het θ -oppervlak. Met behulp van deze θ 's werden de overige vier parameters geplot in figuur 10. De variabiliteit van de b_i 's lijkt relatief beperkt, de α_i 's daarentegen blijken minder robuust. Bovendien vertoont ook deze methode hetzelfde probleem voor relatief grote imaginaire z_i : figuur 11 toont analoog als in sectie 3.2 een contourplot met een doorsnede afgeleid uit de vergelijking ter bepaling van θ , met gekozen verhouding $\alpha = 2.25$. De ontstane cosinussen in de noemers van vergelijking (25) bemoeilijken dus opnieuw de bepaling van θ . Om het effect van een grotere α te testen, werd hetzelfde contourplot met doorsnede gemaakt in figuur 12 voor $\alpha = 25$. Hoewel de positie van de problematische polen niet onderhevig lijkt aan de verhouding, liggen de poollijnen wel steeds dichter bij de nulpunten, dus er dient nog voorzichtiger omgegaan te worden met de numerieke benadering om nog een correct resultaat te krijgen. Samengevat lijkt het dus wenselijk om de methode opnieuw slechts toe te passen op een koppel relatief kleine frequenties.



Figuur 10: Het verloop van de verschillende coëfficiënten in functie van de gefitte parameters.



Figuur 11: Contourplot uit vergelijking (25) voor $\alpha = 2.25$.



Figuur 12: Contourplot uit vergelijking (25) voor $\alpha = 25$.

Voor heel kleine waarden van z_i zijn de parameters opnieuw onderhevig aan afrondingsfouten. Om dit op te lossen kunnen de volgende reeksontwikkelingen gebruikt worden:

$$\theta = \frac{\sqrt{15}}{10} + \frac{\sqrt{15}}{21000}(5z_1^2 + z_2^2) - \frac{\sqrt{15}}{1058400000}(2295z_1^4 + 85z_1^2z_2^2 + 131z_2^4) + \frac{\sqrt{15}}{97796160000000}(1730250z_1^6 - 1653665z_1^4z_2^2 - 5765z_1^2z_2^4 + 26974z_2^6) + \dots \quad (26)$$

$$b_1 = \frac{5}{18} - \frac{z_1^2 + z_2^2}{900} - \frac{48323z_1^2z_2^2}{680400000} + \frac{11401z_2^4}{680400000} + \frac{13908786221z_1^2z_2^4}{123451776000000000} + \frac{13908786221z_1^4z_2^2}{123451776000000000} + \frac{888791z_2^6}{2041200000000} - \frac{888791z_1^6}{2041200000000} + \dots \quad (27)$$

$$b_2 = \frac{4}{9} + \frac{z_1^2 + z_2^2}{450} - \frac{11401z_1^4}{340200000} + \frac{48323z_1^2z_2^2}{340200000} + \frac{37529453779z_1^4z_2^2}{6172588800000000} - \frac{11401z_2^4}{340200000} - \frac{13908786221z_1^2z_2^4}{6172588800000000} + \frac{888791z_2^6}{1020600000000} + \dots \quad (28)$$

$$\alpha_2 = \frac{\sqrt{15}}{15} + \frac{17\sqrt{15}z_1^2}{90000} - \frac{13873\sqrt{15}z_1^4}{1620000000} + \frac{7\sqrt{15}z_2^2}{15000} + \frac{51329\sqrt{15}z_1^2z_2^2}{1620000000} + \frac{19679796547\sqrt{15}z_1^4z_2^2}{293932800000000000} - \frac{9673\sqrt{15}z_2^4}{1620000000} - \frac{10021024253\sqrt{15}z_1^2z_2^4}{293932800000000000} + \frac{5240441\sqrt{15}z_2^6}{340200000000000} + \dots \quad (29)$$

$$\alpha_3 = \frac{\sqrt{15}}{30} + \frac{\sqrt{15}z_1^2}{90000} - \frac{6197\sqrt{15}z_1^4}{2835000000} - \frac{\sqrt{15}z_2^2}{3750} - \frac{1843\sqrt{15}z_1^2z_2^2}{354375000} + \frac{1908300773\sqrt{15}z_1^4z_2^2}{514382400000000000} + \frac{5039\sqrt{15}z_2^4}{1417500000} + \frac{4589757173\sqrt{15}z_1^2z_2^4}{514382400000000000} - \frac{1560841\sqrt{15}z_2^6}{170100000000000} + \dots \quad (30)$$

4.3 Foutterm

De methode is niet langer gemodificeerd, dus voor het onderzoeken van de foutterm kan uitdrukking (7) gebruikt worden en de orde volgt uit lemma 7. De minimale orde van de

h 's waarvoor een ordeterm niet nul is, blijkt bij de drietrapsmethode orde zeven te zijn. Net zoals bij de tweetrapsmethode worden hieronder per orde alle bomen samen met hun bijdrage tot de leidende foutterm opgelijst:

- Bomen van orde één: zesde-ordeterminen in h

$$- \tau \Rightarrow -\frac{h^7}{504000} \mu_1^4 \mu_2^2 f$$

- Bomen van orde drie: vierde-ordeterminen in h

$$- [\tau^2] \Rightarrow -\frac{h^7}{403200} (\mu_2^2 - 2\mu_1^2) \mu_1^2 \{f^2\}$$

$$- [2\tau]_2 \Rightarrow \frac{h^7}{201600} (\mu_2^2 - 2\mu_1^2) \mu_1^2 \{2f\}_2$$

- Bomen van orde vijf: tweede-ordeterminen in h

$$- [\tau^4] \Rightarrow -\frac{h^7}{2016000} (\mu_2^2 + 5\mu_1^2) \{f^4\}$$

$$- [[\tau]^2] \Rightarrow \frac{h^7}{2016000} (2\mu_1^2 - 3\mu_2^2) \{\{f\}^2\}$$

$$- [\tau[\tau^2]] \Rightarrow \frac{h^7}{2016000} (6\mu_2^2 - 5\mu_1^2) \{f\{f^2\}\}$$

$$- [\tau[2\tau]_2] \Rightarrow \frac{h^7}{2016000} (6\mu_2^2 - 40\mu_1^2) \{f\{2f\}_2\}$$

$$- [\tau^2[\tau]] \Rightarrow \frac{h^7}{2016000} (5\mu_1^2 - 6\mu_2^2) \{f^2\{f\}\}$$

$$- [2\tau^3]_2 \Rightarrow \frac{h^7}{2016000} (4\mu_2^2 + 20\mu_1^2) \{2f^3\}_2$$

$$- [2\tau[\tau]]_2 \Rightarrow \frac{h^7}{2016000} (12\mu_2^2 - 10\mu_1^2) \{2f\{f\}\}_2$$

$$- [3\tau^2]_3 \Rightarrow \frac{h^7}{2016000} (5\mu_1^2 - 6\mu_2^2) \{3f\}_3$$

$$- [4\tau]_4 \Rightarrow \frac{h^7}{2016000} (40\mu_1^2 - 6\mu_2^2) \{4f\}_4$$

- Bomen van orde zeven: constante termen

$$- [\tau^6] \Rightarrow \frac{h^7}{2016000} \{f^6\}$$

$$- [[\tau^2]^2] \Rightarrow \frac{h^7}{201600} \{\{f^2\}^2\}$$

$$- [[\tau]^3] \Rightarrow \frac{h^7}{134400} \{\{f\}^3\}$$

$$- [[2\tau]_2^2] \Rightarrow \frac{h^7}{201600} \{\{2f\}_2^2\}$$

$$- [\tau[\tau^4]] \Rightarrow -\frac{h^7}{134400} \{f\{f^4\}\}$$

$$- [\tau[[\tau]^2]] \Rightarrow -\frac{h^7}{44800} \{f\{\{f\}^2\}\}$$

$$- [\tau[\tau[\tau^2]]] \Rightarrow -\frac{h^7}{33600} \{f\{f\{f^2\}\}\}$$

$$- [\tau[\tau][2\tau]_2] \Rightarrow -\frac{h^7}{33600} \{f\{f\{2f\}_2\}\}$$

$$- [\tau[\tau^2[\tau]]] \Rightarrow -\frac{h^7}{22400} \{f\{f^2\{f\}\}\}$$

$$- [\tau[2\tau^3]_2] \Rightarrow \frac{h^7}{100800} \{f\{2f^3\}_2\}$$

$$- [\tau[2\tau[\tau]]_2] \Rightarrow \frac{h^7}{33600} \{f\{2f\{f\}\}_2\}$$

$$\begin{aligned}
- [\tau^2[{}_3\tau^2]_3] &\Rightarrow \frac{h^7}{100800} \{f^2\{{}_3f^2\}_3\} \\
- [\tau[{}_4\tau]_4] &\Rightarrow \frac{h^7}{100800} \{f\{{}_4f\}_4\} \\
- [\tau^2[\tau^3]] &\Rightarrow -\frac{h^7}{100800} \{f^2\{f^3\}\} \\
- [\tau^2[\tau[\tau]]] &\Rightarrow -\frac{h^7}{33600} \{f^2\{f\{f\}\}\} \\
- [\tau^2[\tau]^2] &\Rightarrow \frac{h^7}{44800} \{f^2\{f\}^2\} \\
- [\tau^2[{}_2\tau^2]_2] &\Rightarrow -\frac{h^7}{100800} \{f^2\{{}_2f^2\}_2\} \\
- [\tau^2[{}_3\tau]_3] &\Rightarrow -\frac{h^7}{100800} \{f^2\{{}_3f\}_3\} \\
- [\tau^3[\tau^2]] &\Rightarrow \frac{h^7}{100800} \{f^3\{f^2\}\} \\
- [\tau^3[{}_2\tau]_2] &\Rightarrow \frac{h^7}{100800} \{f^3\{{}_2f\}_2\} \\
- [\tau^4[\tau]] &\Rightarrow \frac{h^7}{134400} \{f^4\{f\}\} \\
- [[\tau^2][{}_2\tau]_2] &\Rightarrow \frac{h^7}{100800} \{\{f^2\}\{{}_2f\}_2\} \\
- [[\tau^3][\tau]] &\Rightarrow -\frac{h^7}{100800} \{\{f^3\}\{f\}\} \\
- [[\tau[\tau]][\tau]] &\Rightarrow -\frac{h^7}{33600} \{\{f\{f\}\}\{f\}\} \\
- [[{}_2\tau^2]_2[\tau]] &\Rightarrow -\frac{h^7}{100800} \{\{{}_2f^2\}_2\{f\}\} \\
- [[{}_3\tau]_3[\tau]] &\Rightarrow -\frac{h^7}{100800} \{\{{}_3f\}_3\{f\}\} \\
- [\tau[\tau][\tau^2]] &\Rightarrow \frac{h^7}{33600} \{f\{f\}\{f^2\}\} \\
- [\tau[\tau][{}_2\tau]_2] &\Rightarrow \frac{h^7}{33600} \{f\{f\}\{{}_2f\}_2\} \\
- [{}_2\tau^5]_2 &\Rightarrow -\frac{h^7}{33600} \{f^5\} \\
- [{}_2\tau[\tau^3]]_2 &\Rightarrow \frac{h^7}{50400} \{{}_2f\{f^3\}\}_2 \\
- [{}_2\tau[\tau[\tau]]]_2 &\Rightarrow \frac{h^7}{16800} \{{}_2f\{f\{f\}\}\}_2 \\
- [{}_2\tau[\tau]^2]_2 &\Rightarrow -\frac{h^7}{22400} \{{}_2f\{f\}^2\}_2 \\
- [{}_2\tau[{}_2\tau^2]_2]_2 &\Rightarrow \frac{h^7}{50400} \{{}_2f\{{}_2f^2\}_2\}_2 \\
- [{}_2\tau[{}_3\tau]_3]_2 &\Rightarrow \frac{h^7}{50400} \{{}_2f\{{}_3f\}_3\}_2 \\
- [{}_2\tau^2[\tau^2]]_2 &\Rightarrow -\frac{h^7}{33600} \{{}_2f^2\{f^2\}\}_2 \\
- [{}_2\tau^2[{}_2\tau]_2]_2 &\Rightarrow -\frac{h^7}{33600} \{{}_2f^2\{{}_2f\}_2\}_2 \\
- [{}_2\tau^3[\tau]]_2 &\Rightarrow -\frac{h^7}{33600} \{{}_2f^3\{f\}\}_2 \\
- [{}_2[\tau][\tau^2]]_2 &\Rightarrow -\frac{h^7}{33600} \{{}_2\{f\}\{f^2\}\}_2 \\
- [{}_2[\tau][{}_2\tau]_2]_2 &\Rightarrow -\frac{h^7}{33600} \{{}_2\{f\}\{{}_2f\}_2\}_2 \\
- [{}_3\tau^4]_3 &\Rightarrow \frac{h^7}{134400} \{{}_3f^4\}_3 \\
- [{}_3[\tau]^2]_3 &\Rightarrow \frac{h^7}{44800} \{{}_3\{f\}^2\}_3 \\
- [{}_3\tau[\tau^2]]_3 &\Rightarrow \frac{h^7}{33600} \{{}_3f\{f^2\}\}_3
\end{aligned}$$

$$\begin{aligned}
- [3\tau[2\tau]_2]_3 &\Rightarrow \frac{h^7}{33600} \{3f\{2f\}_2\}_3 \\
- [3\tau^2[\tau]]_3 &\Rightarrow \frac{h^7}{22400} \{3f^2\{f\}\}_3 \\
- [4\tau^3]_4 &\Rightarrow -\frac{h^7}{100800} \{4f^3\}_4 \\
- [4\tau[\tau]]_4 &\Rightarrow -\frac{h^7}{33600} \{4f\{f\}\}_4 \\
- [5\tau^2]_5 &\Rightarrow -\frac{h^7}{100800} \{5f^2\}_5 \\
- [6\tau]_6 &\Rightarrow -\frac{h^7}{100800} \{6f\}_6
\end{aligned}$$

Het is nu zeer moeilijk om de juiste lineaire combinaties te vinden om de leidende foutterm te herschrijven.

4.4 Stabiliteit

De stabiliteitsfunctie wordt opnieuw gevonden via (10):

$$R(z) = \frac{-(2b_1\alpha_2 - \alpha_3b_2)^2z^3 - 2(2\alpha_2^2b_1 + \alpha_3^2b_2)z^2 - (2b_1b_2 + b_2^2)z - 2b_2}{(2b_1\alpha_2 - \alpha_3b_2)^2z^3 - 2(2\alpha_2^2b_1 + \alpha_3^2b_2)z^2 + (2b_1b_2 + b_2^2)z - 2b_2}. \quad (31)$$

Om opnieuw een expliciete oplossing af te leiden voor het gebied van absolute stabiliteit, wordt de volgende vereenvoudigde notatie ingevoerd:

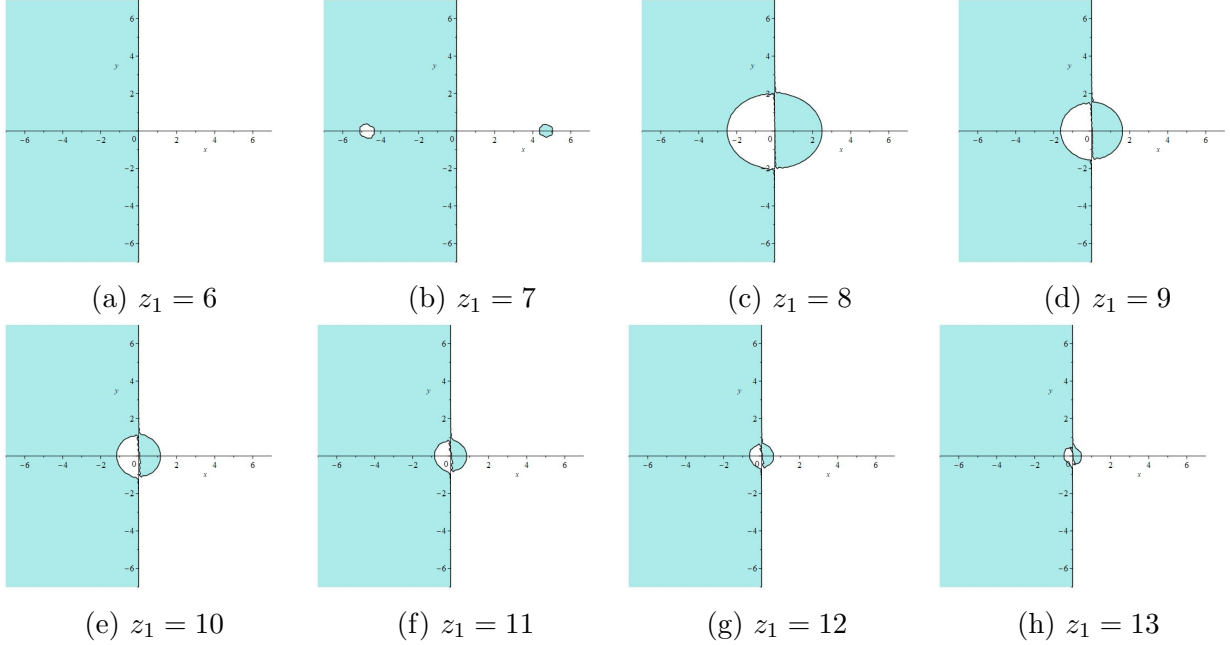
$$R(z) = \frac{-az^3 - bz^2 - cz - d}{az^3 - bz^2 + cz - d}. \quad (32)$$

met dus $a = (2b_1\alpha_2 - \alpha_3b_2)^2$, $b = 2(2\alpha_2^2b_1 + \alpha_3^2b_2)$, $c = (2b_1b_2 + b_2^2)$ en $d = 2b_2$. Hieruit volgt nu de bepaling van de grens van absolute stabiliteit (met $z = x + iy$):

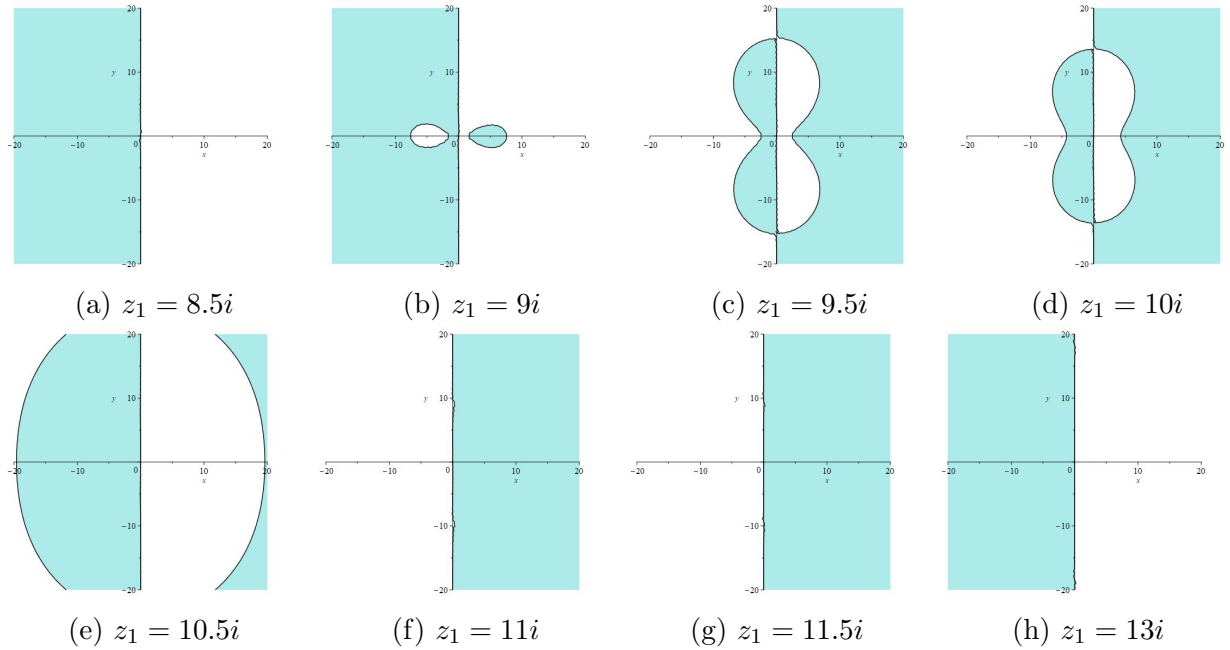
$$\begin{aligned}
&|R(z)| = 1 \\
&\Leftrightarrow |az^3 + bz^2 + cz + d| = |az^3 - bz^2 + cz - d| \\
&\Leftrightarrow |az^3 + bz^2 + cz + d|^2 = |az^3 - bz^2 + cz - d|^2 \\
&\Leftrightarrow |ax^3 + (3ayi + b)x^2 + (2byi - 3ay^2 + c)x - ay^3i + cyi - by^2 + d|^2 \\
&\quad = |ax^3 + (3ayi - b)x^2 + (-2byi - 3ay^2 + c)x - ay^3i + cyi + by^2 - d|^2 \\
&\Leftrightarrow |ax^3 - 3axy^2 + bx^2 - by^2 + cx + d + (3ax^2y - ay^3 + 2bxy + cy)i|^2 \\
&\quad = |ax^3 - 3axy^2 - bx^2 + by^2 + cx - d + (3ax^2y - ay^3 - 2bxy + cy)i|^2 \\
&\Leftrightarrow (ax^3 - 3axy^2 + bx^2 - by^2 + cx + d)^2 + (3ax^2y - ay^3 + 2bxy + cy)^2 \\
&\quad = (ax^3 - 3axy^2 - bx^2 + by^2 + cx - d)^2 + (3ax^2y - ay^3 - 2bxy + cy)^2 \\
&\Leftrightarrow x((x^2 + y^2)^2b + d(x^2 - 3y^2))a + c((x^2 + y^2)b + d) = 0
\end{aligned}$$

Naast de oplossing $x = 0$ bevat de tweede factor nog vier nulpunten in functie van a, b, c en d . De expliciete vorm van deze oplossingen is niet inzichtelijk, maar met numerieke voorbeelden voor de z_i kan getest worden of ze inderdaad werkelijke (of dus reële) nulpunten zijn. Er blijkt dan dat voor bepaalde parameterwaarden, zowel reële als complexe, de oplossing $x = 0$ niet meer de enige oplossing is, of dus dat niet voor alle mogelijke frequenties exacte A-stabiliteit wordt verkregen. In figuur 13 wordt dit geïllustreerd voor reële z_1 : vanaf $z_1 = 7$

ontstaan twee ‘gaten’ in de stabiliteit, die voor grotere z_1 samensmelten en inkrimpen. In figuur 14 werd het imaginaire geval bekeken: het stabiliteitsgebied slaat twee keer om naar het andere halfvlak en voor bepaalde frequenties ontstaan ook ‘gaten’ rond de oorsprong. In beide gevallen werd $\theta = \frac{\sqrt{15}}{10}$ gesteld, waardoor z_2 en dus μ_2 impliciet vastligt.

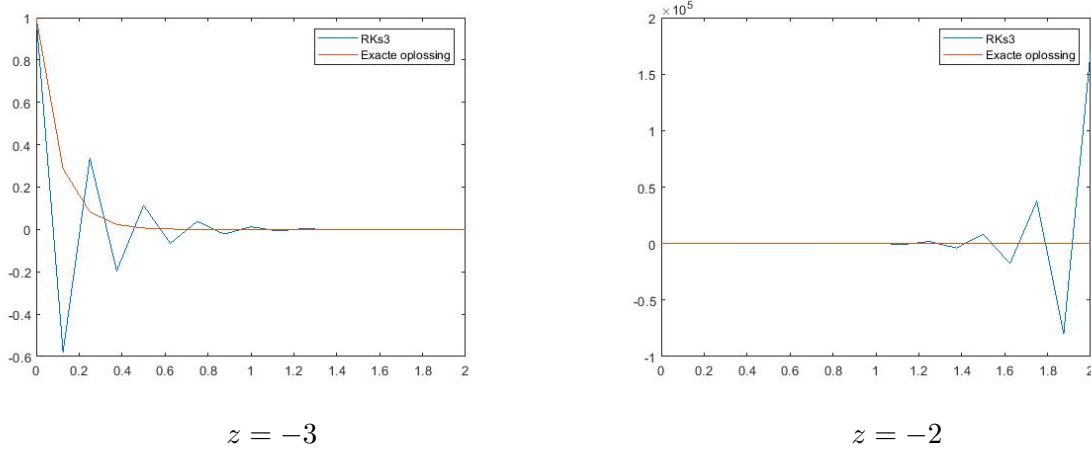


Figuur 13: Stabiliteitsgebieden (turquoise) voor verschillende reële waarden van z_1 .



Figuur 14: Stabiliteitsgebieden (turquoise) voor verschillende imaginaire waarden van z_1 .

Voorbeeld 16. Uit figuur 13c valt af te lezen dat, indien RKs3 gefit wordt op $z_1 = 8$, de testfrequentie $z = \omega h = -3$ binnen het stabiliteitsgebied valt, maar $z = -2$ niet meer. Dit wordt geïllustreerd in figuur 15 over het interval $[0, 2]$: beide exacte oplossingen gaan naar 0, maar de numerieke oplossing bij $z = -2$ explodeert onmiddellijk.



Figuur 15: Numerieke benaderingen binnen en buiten het stabiliteitsgebied.

4.5 Speciale gevallen

Ook bij het drietrapsgeval zijn al enkele specifieke keuzes van z_2 onderzocht. Het geval van gelijke frequenties $\mu_1 = \mu_2$ kan gevonden worden door de limiet te nemen van het linkerlid van (25):

$$G(z_1, z_1) = \lim_{z_2 \rightarrow z_1} G(z_1, z_2) = \frac{\cosh(\frac{z_1}{2}) - \frac{2 \sinh(\frac{z_1}{2})}{z_1}}{z_1 \theta \sinh(z_1 \theta)}.$$

De parameter θ volgt dan uit $G(z_1, z_1) = G(z_1, 2z_1)$, of dus:

$$\frac{\cosh(\frac{z_1}{2}) - \frac{2 \sinh(\frac{z_1}{2})}{z_1}}{z_1 \theta \sinh(z_1 \theta)} = \frac{\frac{\sinh(\frac{z_1}{2})}{\frac{z_1}{2}} - \frac{\sinh(z_1)}{z_1}}{\cosh(z_1 \theta) - \cosh(2z_1 \theta)}. \quad (33)$$

Dit levert helaas geen expliciete oplossing. Een verdere bespreking kan gevonden worden in [16], waarin ook de tweede frequentie wordt ingeruild voor een hogere polynomiale orde, corresponderend met $z_2 = 0$. Dit geeft een eenvoudiger, doch opnieuw onoplosbare vergelijking voor θ :

$$\theta^2 = \frac{z_1(1 - \cosh(\theta z_1))}{12(z_1 - 2 \sinh(\frac{z_1}{2}))}. \quad (34)$$

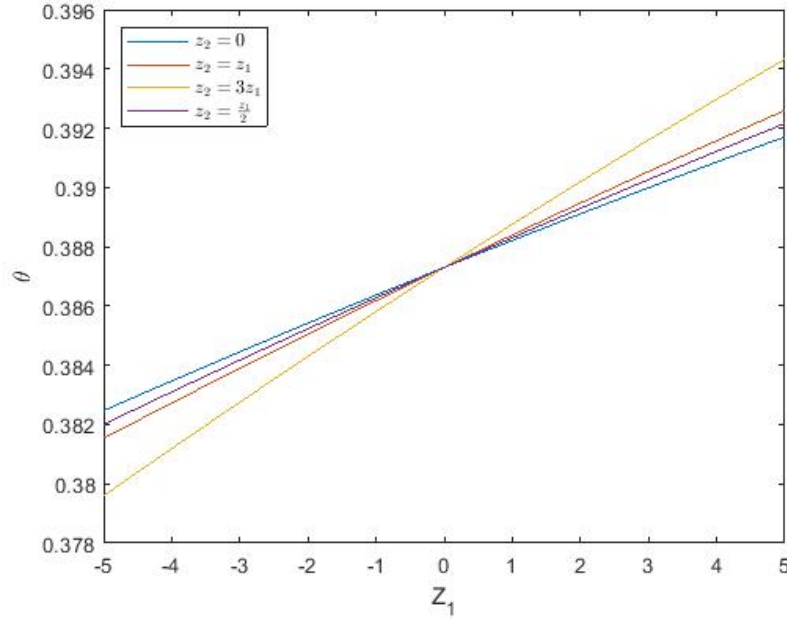
In [3], wordt het geval $z_2 = 3z_1$ besproken, met de volgende exacte oplossing voor θ :

$$\theta = \frac{2}{z_1} \arccos \left(\frac{1}{6} \sqrt{15 + 6 \cosh\left(\frac{z_1}{2}\right)} + 3 \sqrt{15 + 8 \cosh\left(\frac{z_1}{2}\right) + 2 \cosh(z_1)} \right). \quad (35)$$

Een laatste mogelijkheid die een aparte vermelding waard is, is het geval $z_2 = \frac{z_1}{2}$:

$$\theta = \frac{4}{z_1} \arccos \left(\frac{1}{4} \sqrt{6 + 2 \sqrt{9 + 8 \cosh \left(\frac{z_1}{4} \right)^2 + 8 \cosh \left(\frac{z_1}{4} \right)}} \right). \quad (36)$$

Het verloop van θ in deze vier gevallen wordt weergegeven in figuur 16, waar te zien is dat een grotere verhouding $|\alpha|$ zorgt voor een variabelere θ .



Figuur 16: Het verloop van θ voor enkele speciale gevallen van z_2 .

5 Viertrapsmethode

5.1 De methode

Wanneer de symmetrie- en symplecticiteitsvoorwaarden worden toegepast op een standaard viertrapsmethode, wordt het volgende tableau verkregen:

$$\begin{array}{c|c|cccc}
 \frac{1}{2} - \theta_1 & \gamma_1 & \frac{\gamma_1 b_1}{2} & \frac{\gamma_1 b_2}{2} - \alpha_1 & \frac{\gamma_1 b_2}{2} - \alpha_2 & \frac{\gamma_1 b_1}{2} - \alpha_3 \\
 \frac{1}{2} - \theta_2 & \gamma_2 & \frac{\gamma_2 b_1}{2} - \frac{b_1 \alpha_1 \gamma_2}{\gamma_1 b_2} & \frac{\gamma_2 b_2}{2} & \frac{\gamma_2 b_2}{2} - \alpha_4 & \frac{\gamma_2 b_1}{2} + \frac{b_1 \alpha_2 \gamma_2}{\gamma_1 b_2} \\
 \frac{1}{2} + \theta_2 & \gamma_2 & \frac{\gamma_2 b_1}{2} - \frac{b_1 \alpha_2 \gamma_2}{\gamma_1 b_2} & \frac{\gamma_2 b_2}{2} + \alpha_4 & \frac{\gamma_2 b_2}{2} & \frac{\gamma_2 b_1}{2} + \frac{b_1 \alpha_1 \gamma_2}{\gamma_1 b_2} \\
 \frac{1}{2} + \theta_1 & \gamma_1 & \frac{\gamma_1 b_1}{2} + \alpha_3 & \frac{\gamma_1 b_2}{2} + \alpha_2 & \frac{\gamma_1 b_2}{2} + \alpha_1 & \frac{\gamma_1 b_1}{2} \\
 \hline
 & & b_1 & b_2 & b_2 & b_1
 \end{array} \quad (37)$$

Het blijkt ditmaal echter niet zo eenvoudig om stapsgewijs voorwaarden op te leggen aan de interne en externe trappen van dit systeem om tot een oplossing te komen voor alle parameters. Daarom werd ditmaal gestart van een schema met als enige voorwaarde dat er symmetrie is bij de integratiepunten:

$$\begin{array}{c|c|cccc}
 \frac{1}{2} - \theta_1 & \gamma_1 & a_{11} & a_{12} & a_{13} & a_{14} \\
 \frac{1}{2} - \theta_2 & \gamma_2 & a_{21} & a_{22} & a_{23} & a_{24} \\
 \frac{1}{2} + \theta_2 & \gamma_3 & a_{31} & a_{32} & a_{33} & a_{34} \\
 \frac{1}{2} + \theta_1 & \gamma_4 & a_{41} & a_{42} & a_{43} & a_{44} \\
 \hline
 & & b_1 & b_2 & b_3 & b_4
 \end{array} \quad (38)$$

Elke interne trap van dit schema heeft, de integratiepunten buiten beschouwing gelaten, vijf vrijheidsgraden, vandaar dat in elk van de trappen de volgende fitting space wordt opgelegd:

$$S_{\text{int}} = \{1, \exp(\pm \mu_1 x), \exp(\pm \mu_2 x)\}.$$

Dit geeft voor elke trap i een exacte oplossing in functie van c_i , bijgevolg zijn de interne coëfficiënten allemaal bepaald in functie van θ_1 en θ_2 . Door het opleggen van de constante functie wordt $\gamma_i = 1$ voor elke trap, maar de uitdrukkingen voor a_{ij} zijn allerminst bevattelijk en worden daarom hier niet opgelijst. Omwille van de consistentie dienen nu in de eerste plaats dezelfde voorwaarden opgelegd te worden in de externe trap:

$$S_{\text{ext}} = \{1, \exp(\pm \mu_1 x), \exp(\pm \mu_2 x)\}.$$

Dit geeft opnieuw een exacte oplossing voor alle b_i :

$$\begin{aligned}
 b_1 = b_4 &= \frac{z_2 \cosh(z_2 \theta_2) \sinh(\frac{z_1}{2}) - z_1 \cosh(z_1 \theta_2) \sinh(\frac{z_2}{2})}{(\cosh(z_2 \theta_2) \cosh(z_1 \theta_1) - \cosh(z_2 \theta_1) \cosh(z_1 \theta_2)) z_1 z_2} \\
 &= \frac{\eta_0(\frac{Z_1}{4}) \eta_{-1}(Z_2 \theta_2^2) - \eta_0(\frac{Z_2}{4}) \eta_{-1}(Z_1 \theta_2^2)}{2(\eta_{-1}(Z_2 \theta_2^2) \eta_{-1}(Z_1 \theta_1^2) - \eta_{-1}(Z_2 \theta_1^2) \eta_{-1}(Z_1 \theta_2^2))}, \\
 b_2 = b_3 &= \frac{z_1 \sinh(\frac{z_2}{2}) \cosh(z_1 \theta_1) - z_2 \cosh(z_2 \theta_1) \sinh(\frac{z_1}{2})}{(\cosh(z_2 \theta_2) \cosh(z_1 \theta_1) - \cosh(z_2 \theta_1) \cosh(z_1 \theta_2)) z_1 z_2} \\
 &= \frac{\eta_0(\frac{Z_1}{4}) \eta_{-1}(Z_2 \theta_1^2) - \eta_0(\frac{Z_2}{4}) \eta_{-1}(Z_1 \theta_1^2)}{2(\eta_{-1}(Z_2 \theta_1^2) \eta_{-1}(Z_1 \theta_2^2) - \eta_{-1}(Z_2 \theta_2^2) \eta_{-1}(Z_1 \theta_1^2))}.
 \end{aligned}$$

Om de resterende twee parameters te bepalen, θ_1 en θ_2 , wordt S_{ext} nu uitgebreid met vier extra voorwaarden:

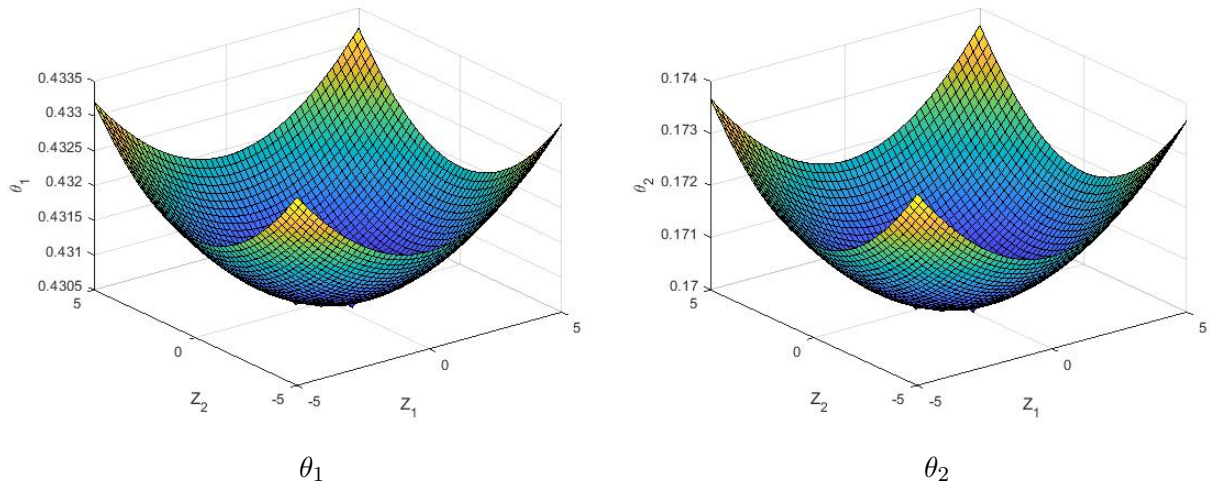
$$S_{\text{ext}} = \{1, x, x^2, x^3, x^4, \exp(\pm\mu_1 x), \exp(\pm\mu_2 x)\}.$$

Dit resulteert in vier vergelijkingen in de theta's, die wegens de gelijkheden $b_1 = b_4$ en $b_2 = b_3$ vereenvoudigd kunnen worden tot twee voorwaarden:

$$\begin{aligned} 2b_1 + 2b_2 - 1 &= 0, \\ 24b_1\theta_1^2 + 24b_2\theta_2^2 - 1 &= 0. \end{aligned} \tag{39}$$

Er bestaat opnieuw geen expliciete oplossing voor de theta's, maar uit dit koppel kunnen θ_1 en θ_2 voor gegeven frequenties μ_1 en μ_2 wel numeriek worden opgelost. Voor $Z_i \in [-5, 5]$ worden de oplossingen geplot in figuur 17. Jammer genoeg blijkt nu dat, door het opleggen van deze voorwaarden, niet voldaan is aan alle symmetrie- en symplecticiteitsvoorwaarden. De verkregen methode ligt dus niet volledig meer in dezelfde lijn als de vorige hoofdstukken.

Notatie 17. In de numerieke resultaten zal voor deze methode de naam *RKs4* gebruikt worden.



Figuur 17: θ 's in functie van Z_1 en Z_2

Opmerking 18. Uiteraard zijn er nog andere mogelijkheden voor de fitting space. Voor de externe trap zijn ook de volgende twee verzamelingen een optie:

$$S_{\text{ext}} = \{1, x, x^2, \exp(\pm\mu_1 x), \exp(\pm\mu_2 x), \exp(\pm\mu_3 x)\},$$

$$S_{\text{ext}} = \{1, \exp(\pm\mu_1 x), \exp(\pm\mu_2 x), \exp(\pm\mu_3 x), \exp(\pm\mu_4 x)\}.$$

In de interne trap zou in plaats van een tweede frequentie ook een hogere polynomiale orde kunnen nagestreefd worden:

$$S_{\text{int}} = \{1, x, x^2, \exp(\pm\mu_1 x)\}.$$

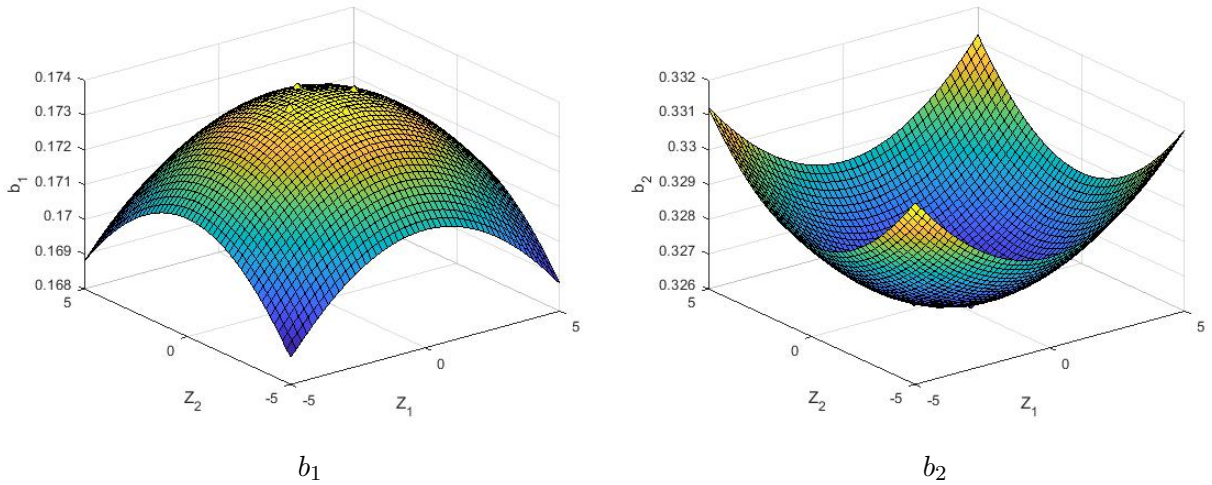
Elk van deze mogelijkheden zou de bepaling en het gebruik van de methode echter nog moeilijker maken: meer frequenties in S_{ext} maken de uitdrukkingen nog lijviger en een frequentie minder in S_{int} bemoeilijkt het zoeken naar de optimale frequenties.

5.2 Coëfficiënten

Een eerste belangrijke observatie is dat de gaussische knooppunten, corresponderend met

$$(\theta_1, \theta_2) = \left(\sqrt{\frac{15 + 2\sqrt{30}}{140}}, \sqrt{\frac{15 - 2\sqrt{30}}{140}} \right),$$

enkel voorkomt bij het geval $(Z_1, Z_2) = (0, 0)$, of dus in het klassieke geval. Daarbuiten lijken de waarden concentrisch rond dit geval gestaag toe te nemen. De b_i -coëfficiënten zijn verder even in beide z_i en wanneer in b_1 de rollen van θ_1 en θ_2 worden omgewisseld, wordt b_2 verkregen en omgekeerd. Dit heeft, samen met de symmetrie van de vergelijkingen in (39), tot gevolg dat de oppervlakken in figuur 17 inderdaad puntsymmetrisch zijn rond de oorsprong en symmetrisch ten opzichte van de assen. De b_i -oppervlakken worden geplott in figuur 18 en hiervoor geldt dezelfde observatie. De overige coëfficiënten worden hier niet weergegeven, maar ook hier zal het oppervlak symmetrisch zijn in de Z_i .



Figuur 18: Het verloop van de b_i 's in functie van de gefitte parameters.

Bij de bepaling van de coëfficiënten is nu nog meer voorzichtigheid nodig: de uitdrukkingen zijn zwaarder, zeker voor de a_{ij} , en het is aangeraden om reeds vanaf een grotere grenswaarde voor $|z_i|$ dan normaal gebruik te maken van reeksontwikkelingen. Voor θ_1 wordt die gegeven door

$$\begin{aligned} \theta_1 = & 0.430568 + 5.51973 \cdot 10^{-5}(z_1^2 + z_2^2) - 7.65209 \cdot 10^{-8}(z_1^4 + z_2^4) - 4.61106 \cdot 10^{-8}z_1^2z_2^2 \\ & + 1.00124 \cdot 10^{-10}(z_1^6 + z_2^6) + 1.77722 \cdot 10^{-10}(z_1^4z_2^2 + z_1^2z_2^4) + \dots \end{aligned} \quad (40)$$

en voor θ_2 door

$$\begin{aligned} \theta_2 = & 0.169990 + 7.45743 \cdot 10^{-5}(z_1^2 + z_2^2) - 1.49694 \cdot 10^{-7}(z_1^4 + z_2^4) + 2.34090 \cdot 10^{-7}z_1^2z_2^2 \\ & + 2.95716 \cdot 10^{-10}(z_1^6 + z_2^6) - 3.87208 \cdot 10^{-10}(z_1^4z_2^2 + z_1^2z_2^4) + \dots \end{aligned} \quad (41)$$

6 Numerieke resultaten

Om de prestaties van elk van de drie methoden te testen zullen vier standaardproblemen bekeken worden waarvoor een exacte oplossing bestaat of waarvan het gedrag al goed gekend is. Telkens zullen de testproblemen numeriek opgelost worden over het interval $[0, 100]$ voor drie verschillende stapgroottes: $h = \frac{1}{2}, \frac{1}{4}$ en $\frac{1}{8}$. Merk op dat deze stapgroottes binair foutloos voor te stellen zijn, zodat eventuele afrondingen in het nemen van stappen al vermeden worden. Vervolgens wordt de verkregen oplossing in elk van de punten vergeleken met de exacte oplossing (of een benadering met machinenauwkeurigheid) en wordt de fout als volgt berekend met de 1-norm:

$$\log_{10} \left[\max_i \|y(x_i) - y_i\|_1 \right].$$

Om ondertussen ook een beeld te krijgen van de optimale frequenties ter bepaling van de parameters in elk van de methoden, wordt heel dit proces gedaan voor een bepaald interval van α 's. Voor elk van de testproblemen bestaat immers al een goeie schatting voor de primaire frequentie μ_1 uit éénparametermethoden, dus de voornaamste open vraag is de bepaling van een tweede frequentie μ_2 , equivalent met de bepaling van een optimale α . In eerste instantie zal steeds gekeken worden naar $\alpha \in [-9 \dots 9]$, maar indien goeie of verwachte resultaten uitblijven, wordt het zoekgebied uitgebreid.

Deze manier van werken heeft bovendien het bijkomende voordeel dat er onmiddellijk een beeld kan gevormd worden over de relatieve performantie ten opzichte van de methoden op één frequentie, zoals afgeleid door Calvo en anderen (zie de secties ‘Speciale gevallen’): deze corresponderen telkens met het geval $\alpha = 0$.

6.1 Het geperturbeerd keplerprobleem

De eerste nuttige test bestaat erin om te kijken of de methode in staat is om een gefitte frequentie effectief op lange termijn correct op te lossen. Hiertoe wordt het geperturbeerde keplerprobleem beschouwd, afkomstig van de volgende hamiltoniaan:

$$H(p, q) = \frac{p_1^2 + p_2^2}{2} - \frac{1}{\sqrt{q_1^2 + q_2^2}} - \frac{\epsilon}{2\sqrt{(q_1^2 + q_2^2)^3}}.$$

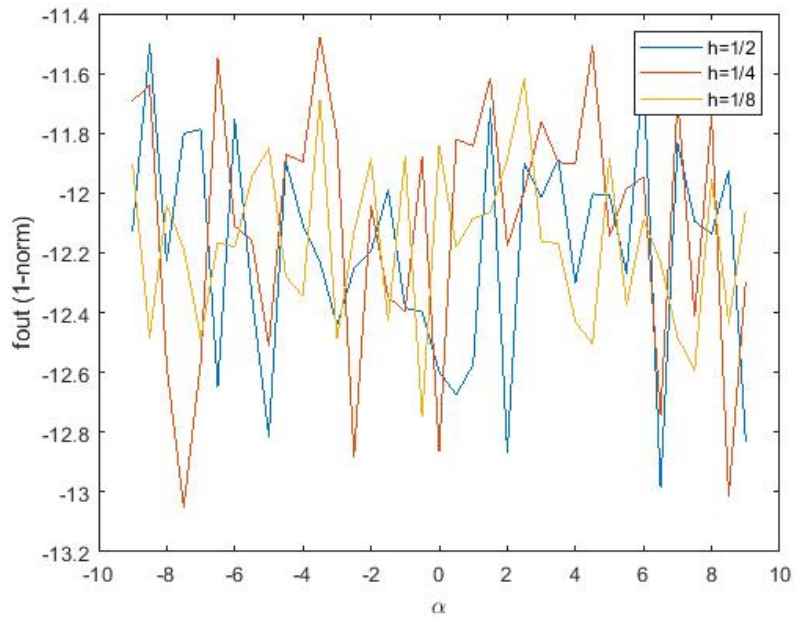
Dit is een versie van het tweelichamenprobleem van Kepler (zie verder), waarin de klassieke potentiële energie wordt gecorrigeerd met een kleine perturbatiefactor ϵ . De hamiltoniaan vertaalt zich in het volgende stelsel (met $(Y_1, Y_2) = (q_1, q_2)$):

$$\begin{cases} Y_1^{(2)} &= -\frac{Y_1}{\sqrt{(Y_1^2 + Y_2^2)^3}} - \frac{Y_1(2\epsilon + \epsilon^2)}{\sqrt{(Y_1^2 + Y_2^2)^5}} \\ Y_2^{(2)} &= -\frac{Y_2}{\sqrt{(Y_1^2 + Y_2^2)^3}} - \frac{Y_2(2\epsilon + \epsilon^2)}{\sqrt{(Y_1^2 + Y_2^2)^5}} \end{cases} \Leftrightarrow \begin{cases} y_1^{(1)} &= y_3 \\ y_2^{(1)} &= y_4 \\ y_3^{(1)} &= -\frac{y_1}{\sqrt{(y_1^2 + y_2^2)^3}} - \frac{y_1(2\epsilon + \epsilon^2)}{\sqrt{(y_1^2 + y_2^2)^5}} \\ y_4^{(1)} &= -\frac{y_2}{\sqrt{(y_1^2 + y_2^2)^3}} - \frac{y_2(2\epsilon + \epsilon^2)}{\sqrt{(y_1^2 + y_2^2)^5}} \end{cases} \quad (42)$$

Dit probleem heeft de exacte oplossing $Y(x) = (-\delta \sin(\delta x), \delta \cos(\delta x))$, met $\delta = 1 + \epsilon$, en zou dus exact moeten worden geïntegreerd wanneer de frequentie μ_1 waarop gefit wordt δ gekozen wordt (of dus $z_1 = i\delta$). In de testen wordt gekozen voor $\epsilon = 0.001$ en uit de exacte oplossing volgt de beginvoorwaarde $y(0) = (1, 0, 0, \delta)$.

6.1.1 Tweetrapsmethode

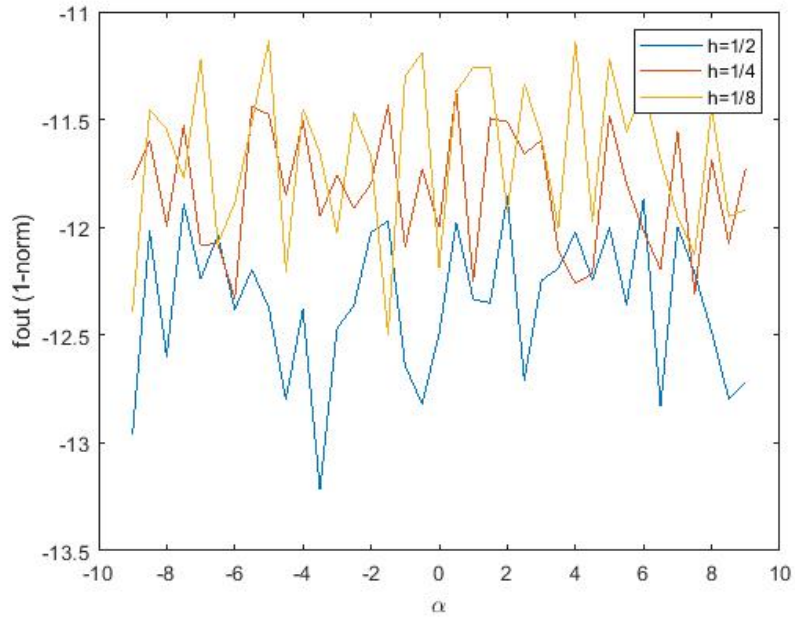
De tweetrapsmethode voldoet exact aan de verwachtingen: in figuur 19 wordt de fout weergegeven. Telkens wordt machinenauwkeurigheid verkregen, slechts onderhevig aan toevallige afrondingsfouten.



Figuur 19: Geperturbeerd keplerprobleem opgelost met RKs2: fout in functie van α

6.1.2 Drietrapsmethode

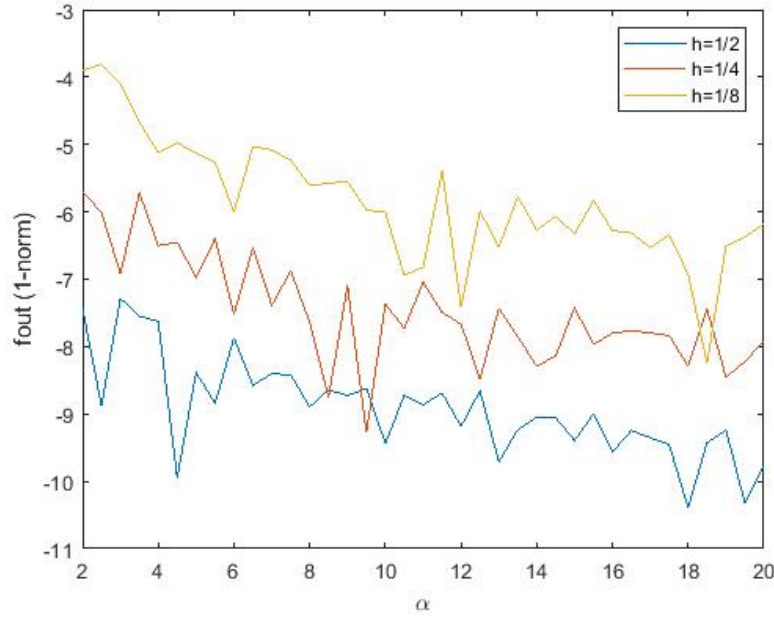
Ook bij het drietrapsgeval wordt bijna machinenauwkeurigheid bereikt voor alle mogelijke α 's, zie figuur 20. Het resultaat is hier wel sterk afhankelijk van de precisie waarmee gewerkt wordt bij het oplossen van de impliciete vergelijkingen bij de bepaling van θ en bij het wegwerken van de interne trappen. De standaardnauwkeurigheid levert povere resultaten, waarbij de fouten zelfs opstapelen bij kleinere h .



Figuur 20: Geperturbeerd keplerprobleem opgelost met RKs3: fout in functie van α

6.1.3 Viertrapsmethode

Bij RKs4 is het niet onmiddellijk duidelijk dat de methode exact is voor het geperturbeerde keplerprobleem. In figuur 21 wordt de fout weergegeven in functie van α . Twee observaties kunnen gemaakt worden: de precisie lijkt af te nemen bij een kleinere stapgrootte, terwijl ze toeneemt bij een grotere verhouding $|\alpha|$ (hetzelfde beeld komt voor bij negatieve α 's). Beide zijn hoogstwaarschijnlijk het gevolg van opstapelende afrondingsfouten. Bij een kleinere h worden beide z_i eveneens kleiner en een kleinere verhouding nadert tot de gevallen $\alpha = 1$ en $\alpha = 0$, waarvoor benaderende reeksontwikkelingen van de coëfficiënten nodig zijn.



Figuur 21: Geperturbeerd keplerprobleem opgelost met RKs4: fout in functie van α

6.2 Een autonoom en lineair probleem

Een logische volgende stap is een systeem met twee verschillende frequenties. Veronderstel het volgende eenvoudige probleem:

$$\begin{cases} Y_1^{(2)} = -k_1^2 Y_1 \\ Y_2^{(2)} = -k_2^2 Y_2 \end{cases} \Leftrightarrow \begin{cases} y_1^{(1)} = y_2 \\ y_2^{(1)} = -k_1^2 y_1 \\ y_3^{(1)} = y_4 \\ y_4^{(1)} = -k_2^2 y_3 \end{cases} \quad (43)$$

De exacte oplossing hiervoor wordt gegeven door $Y_i(x) = \sin(k_i x)$, met dus als beginvoorwaarde $y(0) = (0, k_1, 0, k_2)$.

6.2.1 Tweetrapsmethode

Om dit probleem zo goed mogelijk te integreren kan er op zoek gegaan worden naar de geschikte μ_i zodanig dat de orde van de fout zo hoog mogelijk wordt. De leidende foutterm wordt:

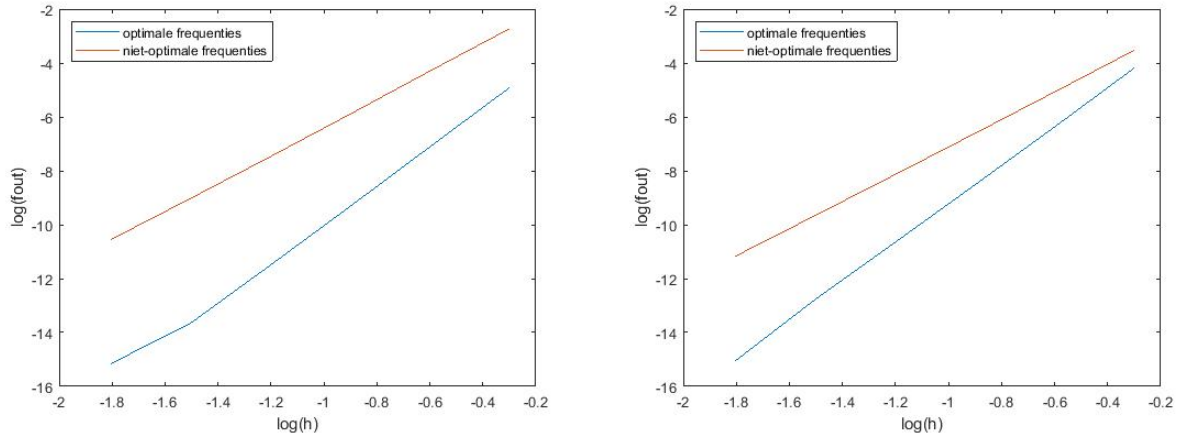
$$\frac{h^5}{2160} J \begin{bmatrix} (3k_1^2 - \mu_2^2 + 6\mu_1^2)(k_1^2 + \mu_1^2) \\ (3k_1^2 - \mu_2^2 + 6\mu_1^2)(k_1^2 + \mu_1^2) \\ (3k_2^2 - \mu_2^2 + 6\mu_1^2)(k_2^2 + \mu_1^2) \\ (3k_2^2 - \mu_2^2 + 6\mu_1^2)(k_2^2 + \mu_1^2) \end{bmatrix}.$$

Deze term wordt nul voor de volgende waarden van μ_i :

$$\left\{ \mu_1 = \pm i k_1, \mu_2 = \pm \sqrt{3k_2^2 - 6k_1^2} \right\}, \quad \left\{ \mu_1 = \pm i k_2, \mu_2 = \pm \sqrt{3k_1^2 - 6k_2^2} \right\}.$$

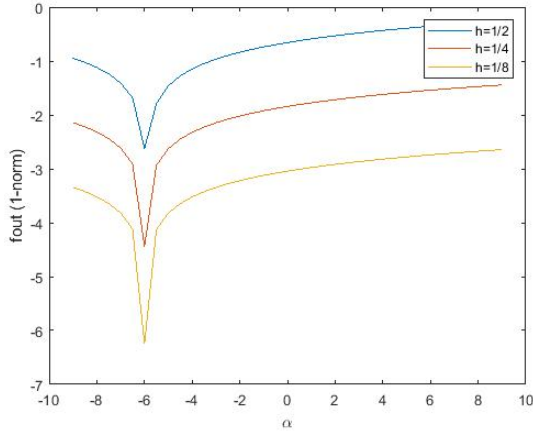
De eerste frequentie van elk koppel komt overeen met een werkelijke frequentie van de oplossing, wat te verwachten is gezien de eerste door de methode telkens exact wordt geïntegreerd. De tweede is een gemengde term van beide werkelijke frequenties. Met deze waarden van μ_i is de orde van de methode dus verhoogd van vier naar minstens vijf. Nu blijkt echter dat de volgende term in de foutontwikkeling dezelfde nulpunten heeft, waardoor de orde van de methode verder verhoogt naar zes.

Dit wordt geïllustreerd in figuur 22, waar een logaritme van de fout wordt weergegeven in functie van de logaritme van de stapgrootte en bijgevolg de richtingscoëfficiënt van de rechte overeenkomt met de orde van de methode. Hierbij werden k_1 en k_2 respectievelijk 1 en 2 gekozen, met voor (μ_1, μ_2) de overeenkomstige correcte waarden $(i, \sqrt{6})$ links en $(2i, \sqrt{21}i)$ rechts, en telkens de niet-optimale waarden $(i, 2i)$ ter referentie. De orde verhoogt links van 4.20 naar 5.91 en rechts zelfs naar 6.16, al is het verschil eerder te wijten aan het feit dat in het linkergeval sneller machinenauwkeurigheid wordt bereikt, waardoor de richtingscoëfficiënt iets afgezwakt wordt. Dit laatste is overigens opmerkelijk, aangezien de tweede frequentie daar verder ligt van de originele en zelfs niet langer complex is.

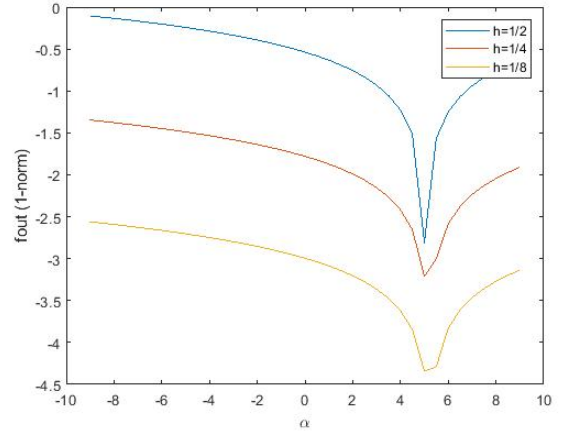


Figuur 22: Log-logdiagrammen voor telkens een optimaal en een niet-optimaal koppel frequenties

Te verwachten is dus dat de geoptimaliseerde frequentie effectief de hoogste nauwkeurigheid geeft. Hiertoe wordt opnieuw het geval $(k_1, k_2) = (1, 2)$ bekeken en wordt de methode tweemaal gefit, voor beide mogelijke primaire frequenties: $\mu_1 = i$ en $\mu_1 = 2i$. De verkregen minima situeren zich inderdaad rond respectievelijk $\alpha = \frac{\sqrt{6}^2}{i^2} = -6$ en $\alpha = \frac{(\sqrt{21}i)^2}{(2i)^2} = 5.25$, zie figuur 23.



(a) $\mu_1 = i$



(b) $\mu_1 = 2i$

Figuur 23: Probleem (43) opgelost met RKs2: fout in functie van α

6.2.2 Drietrapsmethode

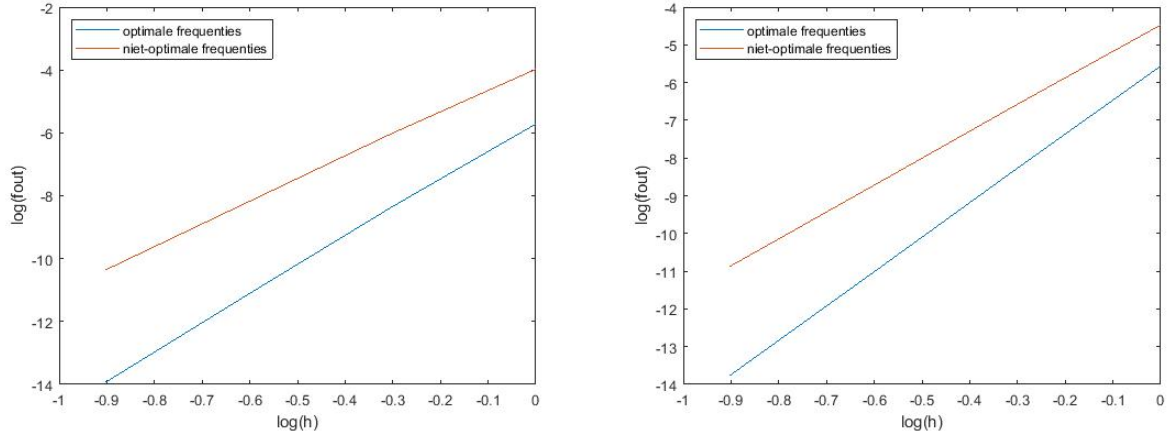
Bij het drietrapsgeval vereenvoudigt de leidende foutterm tot:

$$\frac{h^7}{10080} J \begin{bmatrix} 10(k_1^2 + \mu_1^2)(10k_1^4 - 3k_1^2\mu_2^2 + 10k_1^2\mu_1^2 - 2\mu_2^2\mu_1^2) \\ -(k_1^2 + \mu_1^2)(10k_1^4 - 3\mu_2^2k_1^2 + 10\mu_1^2k_1^2 - 2\mu_2^2\mu_1^2)k_1^2 \\ 10(k_2^2 + \mu_1^2)(10k_2^4 - 3k_2^2\mu_2^2 + 10k_2^2\mu_1^2 - 2\mu_2^2\mu_1^2) \\ -(k_2^2 + \mu_1^2)(10k_2^4 - 3\mu_2^2k_2^2 + 10\mu_1^2k_2^2 - 2\mu_2^2\mu_1^2)k_2^2 \end{bmatrix}.$$

Deze term heeft verschillende nulpunten, zowel zuiver reële, zuiver imaginaire als niet-zuivere. De interpretatie van deze laatste ligt echter niet voor de hand, waardoor enkel de zuivere nulpunten worden bekeken:

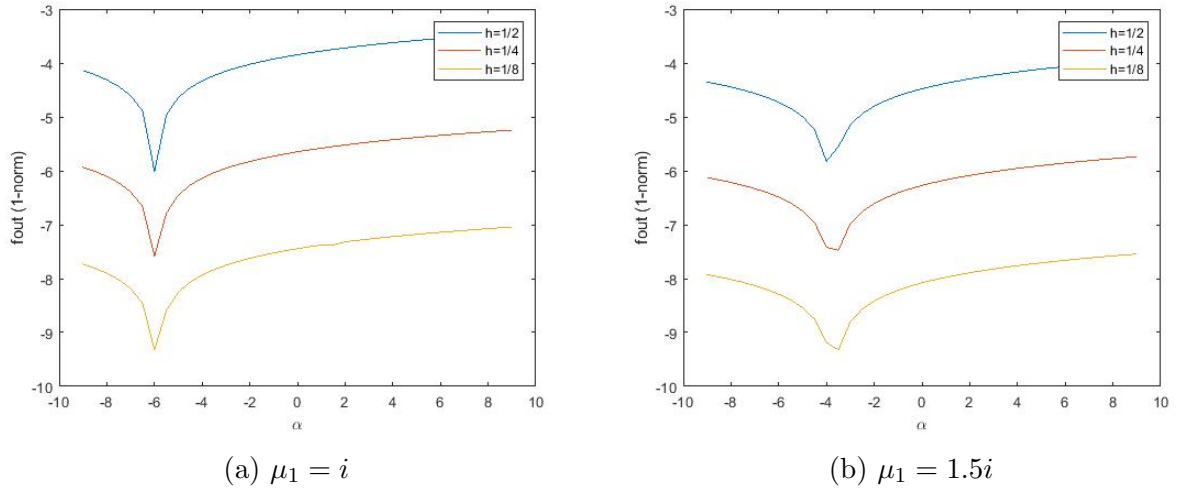
$$\left\{ \mu_1 = \pm ik_1, \mu_2 = \pm \sqrt{\frac{10(k_1^2k_2^2 - k_2^4)}{2k_1^2 - 3k_2^2}} \right\}, \quad \left\{ \mu_1 = \pm ik_2, \mu_2 = \pm \sqrt{\frac{10(k_1^2k_2^2 - k_1^4)}{2k_2^2 - 3k_1^2}} \right\}.$$

Opnieuw komt het eerste nulpunt overeen met een werkelijke frequentie van het systeem en is de tweede een gecorrigeerde frequentie op basis van beide werkelijke μ_i 's, om te compenseren voor de inexacte benaderingen door de interne trappen. Opnieuw blijkt dat er een zekere symmetrie aanwezig is in de termen van de foutontwikkeling, want ook de foutterm bij h^8 verdwijnt met deze nulpunten. Een illustratie van de verhoging van de orde van zes naar acht volgt in figuur 24. Voor de frequenties $(k_1, k_2) = (1, 1.5)$ (anders dan in het tweetrapsgeval, omdat voor het koppel $(z_1, 2z_1)$ een speciale versie van RKs3 wordt verkregen) worden de nulpunten $(i, \sqrt{\frac{225}{38}})$ en $(1.5i, \sqrt{\frac{50}{6}})$, en met deze nulpunten verhoogt de orde telkens van 6.07 naar 8.09. Hierbij werd vergeleken met de orde van het werkelijke koppel frequenties $(i, 1.5i)$.



Figuur 24: Log-logdiagrammen voor telkens een optimaal en een niet-optimaal koppel frequenties

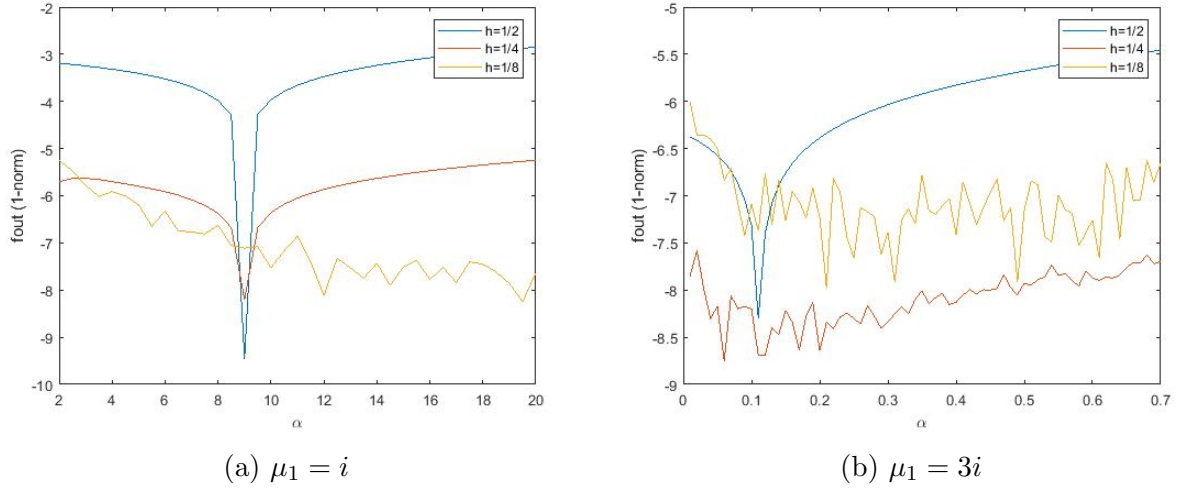
Dezelfde frequenties $(k_1, k_2) = (1, 1.5)$ werden ook gekozen bij het zoeken van een optimale verhouding α . Bij $\mu_1 = i$ zou de optimale verhouding $\alpha = \frac{\sqrt{\frac{225}{38}}^2}{i^2} \approx -5.92$ moeten zijn, bij $\mu_1 = 1.5i$ volgt $\alpha = \frac{\sqrt{\frac{50}{6}}^2}{(1.5i)^2} \approx -3.70$. In figuur 25 is te zien dat voor beide gevallen het verwachte gedrag volgt.



Figuur 25: Probleem (43) opgelost met RKs3: fout in functie van α

6.2.3 Viertrapsmethode

Bij het viertrapsgeval is gekozen voor $(k_1, k_2) = (i, 3i)$, dus de logische keuzes voor de fitting frequenties zijn $(\mu_1, \mu_2) = (i, 3i)$ en $(\mu_1, \mu_2) = (3i, i)$. In figuur 26 worden beide keuzes voor μ_1 getest en de verwachte μ_2 komt inderdaad naar voor: de optima zijn $\alpha = \frac{(3i)^2}{i^2} = 9$ en $\alpha = \frac{i^2}{(3i)^2} = \frac{1}{9}$.



Figuur 26: Probleem (43) opgelost met RKs4: fout in functie van α

Het gedrag van RKs4 bij het geperturbeerd keplerprobleem is ook hier te zien: zodra h en dus de z_i relatief klein worden, zorgen opstapelingen van afrondingsfouten voor een slechter resultaat dan verwacht. In het geval $h = \frac{1}{8}$ komt bijvoorbeeld geen duidelijk minimum meer naar voor en soms presteert een kleinere h zelfs beter.

6.3 Het keplerprobleem

Het tweelichamenprobleem van Kepler bestaat erin de posities en snelheden te vinden van twee lichamen die mekaar aantrekken. De hamiltoniaan van dit probleem is de volgende:

$$H(p, q) = \frac{p_1^2 + p_2^2}{2} - \frac{1}{\sqrt{q_1^2 + q_2^2}}.$$

De vertaling naar een stelsel geeft (stel opnieuw $(Y_1, Y_2) = (q_1, q_2)$):

$$\begin{cases} Y_1^{(2)} \\ Y_2^{(2)} \end{cases} = \begin{cases} -\frac{Y_1}{\sqrt{(Y_1^2 + Y_2^2)^3}} \\ -\frac{Y_2}{\sqrt{(Y_1^2 + Y_2^2)^3}} \end{cases} \Leftrightarrow \begin{cases} y_1^{(1)} \\ y_2^{(1)} \\ y_3^{(1)} \\ y_4^{(1)} \end{cases} = \begin{cases} y_3 \\ y_4 \\ -\frac{y_1}{\sqrt{(y_1^2 + y_2^2)^3}} \\ -\frac{y_2}{\sqrt{(y_1^2 + y_2^2)^3}} \end{cases} \quad (44)$$

De oplossing wordt bekeken met als beginvoorwaarde $y(0) = \left(1 - e, 0, 0, \sqrt{\frac{1+e}{1-e}}\right)$, waarbij e , de excentriciteit van de baan, 0.001 wordt gekozen. Er bestaat geen expliciete oplossing, maar er kan een schatting met machinenaauwkeurigheid gevonden worden uit de volgende impliciete oplossing:

$$y(t) = (\cos(E) - e, \sqrt{1 - e^2} \sin(E), y_1^{(1)}, y_2^{(1)}),$$

met $t = E - e \sin(E)$. Een fourieranalyse van het probleem geeft dat de primaire frequenties van het stelsel $\mu_1 = i$ en $\mu_2 = 0$ zijn.

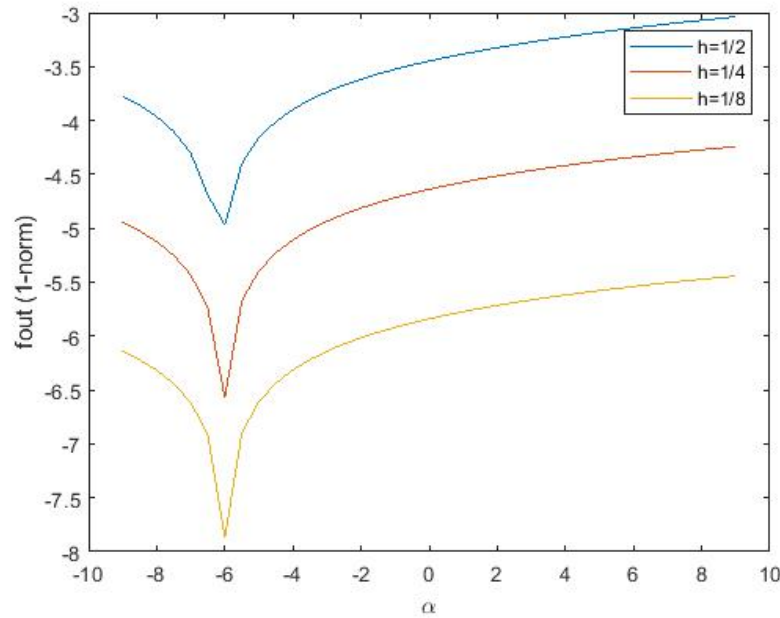
6.3.1 Tweetrapsmethode

De primaire frequentie komt ook naar voor bij het onderzoeken van de foutterm. De leidende foutterm bevat de variabele e en kan er bijgevolg in reeks ontwikkeld rond worden:

$$lte = \frac{h^5}{4320} \begin{bmatrix} 2 \sin(E)(\mu_2^2 \mu_1^2 - 6\mu_1^4 + \mu_2^2 - 9\mu_1^2 - 3) + \mathcal{O}(e) \\ 2(6\mu_1^4 - \mu_2^2 \mu_1^2 + 21) \cos(E) - (\mu_2^2 - 9\mu_1^2 + 18) \cos(E) + \mathcal{O}(e) \\ 2(63\mu_2^2 + 18\mu_1^2) \cos(E) + (24\mu_1^4 + (27 - 7\mu_2^2)\mu_1^2 - 87 - 3\mu_2^2) \cos(E) + \mathcal{O}(e) \\ \sin(E)(36\mu_1^2 - 7\mu_2^2 + 39 + 24\mu_1^4 + (27 - 7\mu_2^2)\mu_1^2) + \mathcal{O}(e) \end{bmatrix}.$$

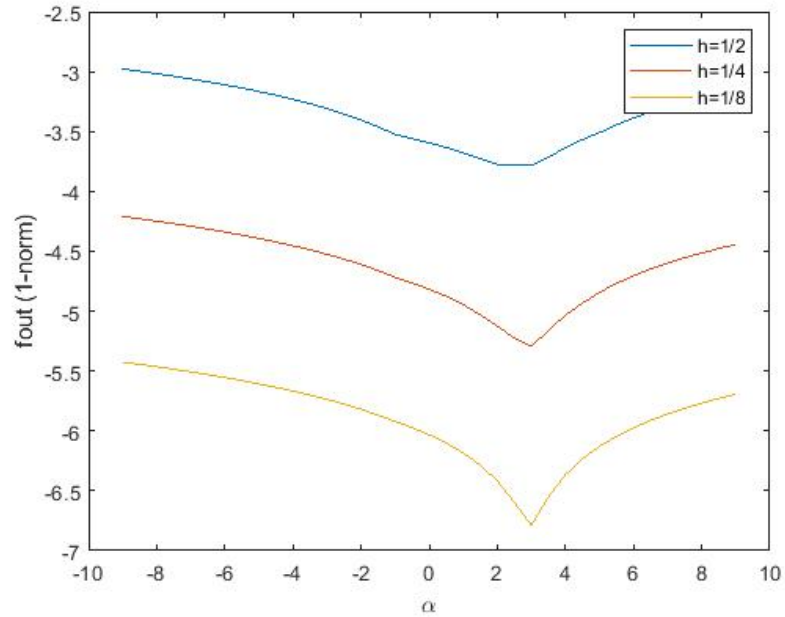
Wanneer enkel wordt gekeken naar de constante termen van deze ontwikkeling (aangezien e klein is, is de rest van de reeks verwaarloosbaar), dan is de frequentie die de term nul maakt inderdaad $\mu_1 = i$. De tweede frequentie volgt hieruit niet en is dus nog vrij te kiezen. Een minimalisatie van de resterende ordetermen in e brengt ook geen duidelijkheid.

De numerieke resultaten bij de schatting van α geven wel meer informatie, zie figuur 27. Er komt geen optimale complexe μ_2 naar voor, maar indien de tweede frequentie reëel wordt gekozen, ontstaat wel een minimum: $\alpha \approx -6$ en dus $\mu_2 \approx 2.45$ lijkt ideaal.



Figuur 27: Keplerprobleem opgelost met RKs2, constante frequenties: fout in functie van α

Men kan ook werken met een variabele frequentie: de vorm van het linkerstelsel in (44) suggereert een frequentie gelijk aan $(Y_1^2 + Y_2^2)^{-\frac{3}{4}}$. Uit de literatuur blijkt echter dat de gecorrigeerde $(Y_1^2 + Y_2^2)^{-\frac{3}{2}}$ beter presteert. In figuur 28 wordt gezocht naar de optimale verhouding met deze versie, waaruit blijkt dat ook daarmee wel degelijk een goeie tweede frequentie kan gekozen worden.

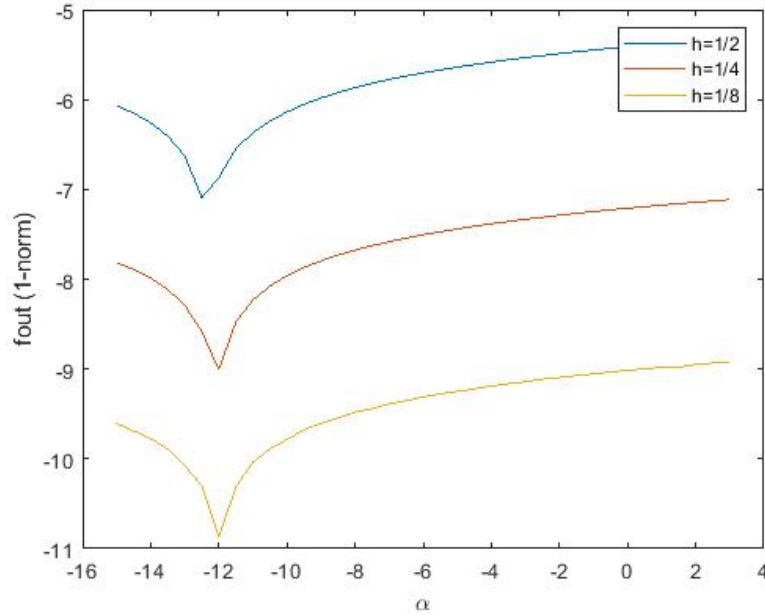


Figuur 28: Keplerprobleem opgelost met RKs2, variabele frequenties: fout in functie van α

Deze aanpak resulteert echter niet in een beter optimum dan het constante geval (in termen van een lagere fout) en bovendien is het computationeel veel zwaarder om bij elke stap de nieuwe frequenties en hiermee de nieuwe coëfficiënten te berekenen. In de volgende methodes zal dit geval dus niet meer beschouwd worden.

6.3.2 Drietrapsmethode

Uit de leidende foutterm volgt nu niet langer informatie over de optimale frequenties, maar omdat μ_1 exact wordt geïntegreerd ligt het voor de hand om daarvoor opnieuw de primaire frequentie van het systeem te kiezen: $\mu_1 = i$. Om de optimale μ_2 te bepalen worden in figuur 29 een reeks tweede frequenties bekeken, met $\alpha \approx -12$ als beste waarde. Hieruit volgt dan dat het optimale koppel $(\mu_1, \mu_2) \approx (i, 3.46)$ is.



Figuur 29: Keplerprobleem opgelost met RKs3: fout in functie van α

6.4 Het eulerprobleem

De koppelvrije versie van Eulers vergelijkingen voor het beschrijven van starre lichamen worden gegeven door:

$$\begin{cases} y_1^{(1)} &= (\alpha - \beta)y_2y_3 \\ y_2^{(1)} &= (1 - \alpha)y_1y_3 \\ y_3^{(1)} &= (\beta - 1)y_1y_2 \end{cases} \quad (45)$$

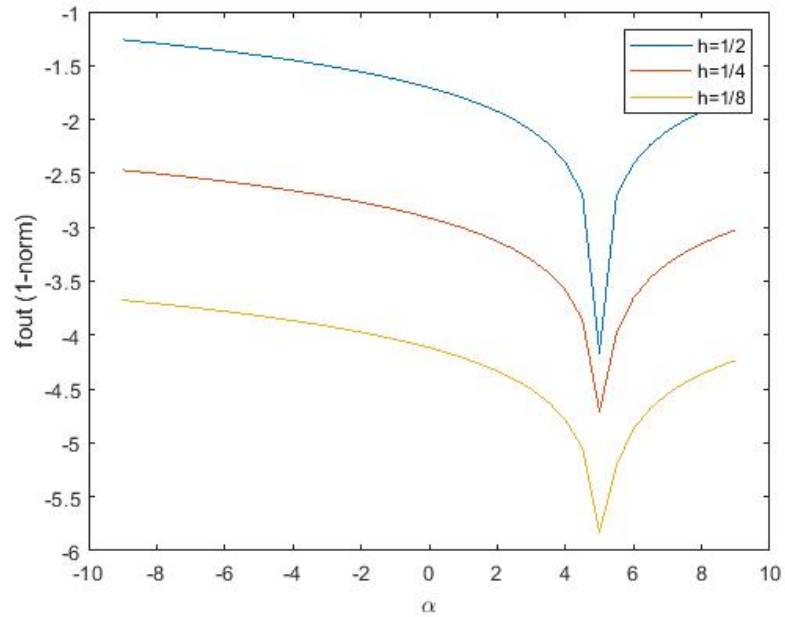
Voor het bepalen van de oplossing werden $\alpha = 1 + \frac{1}{\sqrt{1.51}}$ en $\beta = 1 - \frac{0.51}{\sqrt{1.51}}$ gekozen, met als beginvoorwaarde $y(0) = (0, 1, 1)$. De exacte oplossing wordt dan, in termen van de elliptische Jacobi-functies:

$$y(t) = (\sqrt{1.51} \operatorname{sn}(t, 0.51), \operatorname{cn}(t, 0.51), \operatorname{dn}(t, 0.51)).$$

Met behulp van een fourieranalyse blijken de primaire frequenties van het systeem $\frac{2\pi i}{T} \approx 0.84i$ en 0 te zijn, waarbij $T = 7.45056320933095$.

6.4.1 Tweetrapsmethode

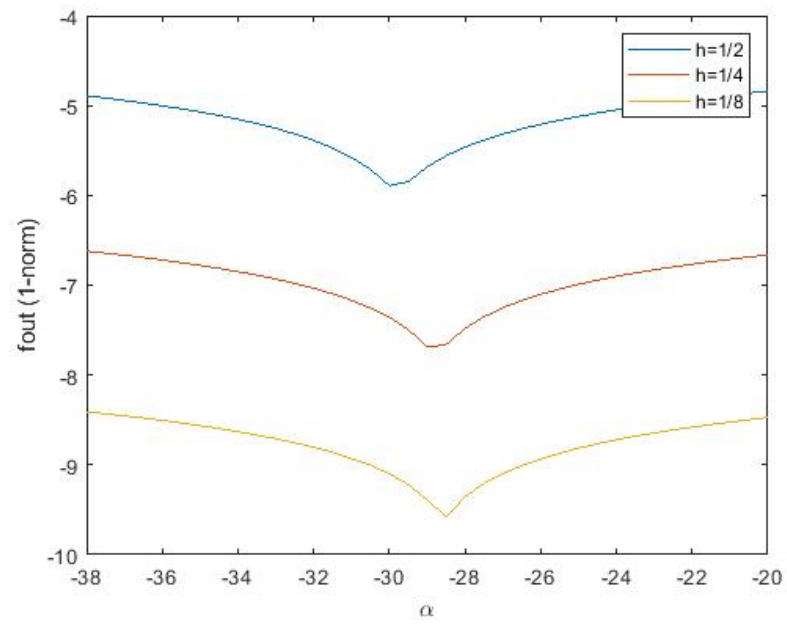
De foutterm geeft hier niet onmiddellijk een indicatie welke frequenties het best gekozen worden, dus wordt voor μ_1 de klassieke waarde $0.84i$ gekozen. In figuur 30 is te zien dat de optimale verhouding α ongeveer 5 is, wat overeen komt met koppel frequenties $(\mu_1, \mu_2) \approx (0.84i, 1.88i)$.



Figuur 30: Eulerprobleem opgelost met RKs2: fout in functie van α

6.4.2 Drietrapsmethode

Ook hier wordt gekozen voor $\mu_1 = 0.84i$. Na het bereik voor α iets op te schuiven, blijkt dat de beste verhouding ongeveer -29 is, te zien in figuur 31. Het minimum is hier echter veel minder uitgesproken. Het corresponderende beste koppel frequenties bevat nu een reële en een complexe waarde: $(\mu_1, \mu_2) \approx (0.84i, 4.52)$.



Figuur 31: Eulerprobleem opgelost met RKs3: fout in functie van α

7 Besluit

Multiparametrische EFRK-methodes lijken op papier een veelbelovend instrument om oscillerende of exponentiële systemen met meerdere niet-verwaarloosbare frequenties goed te integreren. Wanneer echter verder gekeken wordt, vertonen ze enkele belangrijke pijnpunten. In de eerste plaats is er de afleiding, die vaak geen exacte uitdrukkingen voor de coëfficiënten oplevert en computationeel zwaarder en zwaarder wordt naarmate meer trappen en meer frequenties worden toegevoegd. De gebruikte uitdrukkingen worden immers langer en bevatten meer bewerkingen, wat een hogere precisie vereist om op het einde nog het gewenste resultaat te krijgen. Ook op vlak van interpreteerbaarheid geeft dit problemen: het onderzoek naar bijvoorbeeld de stabiliteit en de foutterm is allerm minst eenvoudig.

Een ander groot minpunt ontstaat pas bij het effectief toepassen van de methodes. Er kan wel degelijk precisiewinst gemaakt worden in vergelijking met de éénparametermethodes, maar die is sterk afhankelijk van hoe goed de schattingen zijn van de frequenties die in het systeem voorkomen. Bij de gebruikte testproblemen was inderdaad te zien dat voor een geschikte primaire frequentie de verhouding α telkens een optimum had, maar wel in een zeer beperkt interval. Wanneer een verhouding wordt gekozen die te ver ligt van dat optimum, kunnen zelfs slechtere resultaten uitgekomen worden dan in het éénparametergeval. Wanneer dan een methode wordt gebruikt met een frequentie die niet exact wordt geïntegreerd door de interne trappen en dus niet exact een frequentie van het probleemdomen is (maar een correctie), maakt dat de zaak nog lastiger. Bovendien is voor die correctie een onderzoek van de foutterm nodig, wat ons terugbrengt bij het probleem uit de vorige paragraaf.

Tot slot wordt ook de stabiliteit uit het klassieke gaussische geval niet langer behouden. Er ontstaan al enkele imaginaire probleemfrequenties bij het tweetrapsgeval en bij drie trappen zijn er ook reële z_i met een abnormaal gedrag. Omdat de complexiteit van de stabiliteitsveelterm toeneemt bij meer trappen (en dus meer nulpunten heeft), is de verwachting dat dit ook zo blijft voor hogere s . Er is dus extra oplettendheid nodig bij numerieke oplossingen over grote intervallen.

Appendix

A English summary

Many systems of differential equations have solutions with oscillatory or exponential behaviour. Therefore, when solving these systems, it is beneficial to use numeric solvers that are trained on this behaviour, so-called exponentially fitted methods or EF-methods. They are often designed to exactly integrate a set S consisting of polynomials as well as exponential functions:

$$S_{k,l} = \{1, x, \dots, x^k\} \cup \{\exp(\pm\mu_1 x), \exp(\pm\mu_2 x), \dots, \exp(\pm\mu_l x)\}.$$

In the context of EF-Runge-Kutta-methods (EFRK-methods), many research has been done in the cases where $\mu_1 = \dots = \mu_l$ and $\mu_i = i\mu_1$, but the case of truly unrelated frequencies has been rarely studied. In this masters dissertation, symplectic and symmetric two-, three- and four-stage-EFRK-methods are investigated and an attempt is made to find which frequencies are the best to use, given the frequencies of the problem. Short summaries of the properties and results for each method are listed below.

A.1 Two-stage method

The internal and external stages of the method are exact for the following sets respectively:

$$S_{\text{int}} = \{\exp(\pm\mu_1 x)\}, \quad S_{\text{ext}} = \{1, \exp(\pm\mu_1 x), \exp(\pm\mu_2 x)\}.$$

The coefficients seem quite constant, but in the case of larger imaginary frequencies, the numerical approximation becomes unstable, so the use is best limited to relatively small frequencies. Investigation of the error term learns that the method is of order four, and that the possibility exists to increase to order six, as shown in numerical examples. Unfortunately, this requires very good estimations of the real frequencies of the problem and numerical approximations of the higher order derivatives, which is not realistic for most systems. The stability properties of the method are inherited from its classic gaussian variant for all real frequencies and several complex frequencies, for which the method is exactly-A-stable.

A.2 Three-stage method

The internal and external fitting spaces now become:

$$S_{\text{int}} = \{1, \exp(\pm\mu_1 x)\}, \quad S_{\text{ext}} = \{1, \exp(\pm\mu_1 x), \exp(\pm 2\mu_1 x), \exp(\pm\mu_2 x)\}.$$

The coefficients are slightly more variable than in the two-stage case and the main issue remains: larger imaginary fitting frequencies make the determination of the coefficients difficult. The standard order is six and is adjustable to eight based on the error term, but the expression for the latter is far too difficult in most cases to provide good estimations of the fitting frequencies. The exact-A-stability is now preserved for less frequencies: instability arises too in some real cases.

A.3 Four-stage method

The investigated case made use of the following fitting spaces:

$$S_{\text{int}} = \{1, \exp(\pm\mu_1 x), \exp(\pm\mu_2 x)\}, \quad S_{\text{ext}} = \{1, x, x^2, x^3, x^4, \exp(\pm\mu_1 x), \exp(\pm\mu_2 x)\},$$

although other options are possible. Because the chosen frequencies are now all fitted for both internal and external stages, it is obvious that they must equal the two primary frequencies of the problem, so there is less need to investigate the complicated error term. However, a new issue arises: the coefficients become quite extensive, which makes them more prone to rounding errors. Numerical examples show that the use of even relatively big step sizes can cause biased results.

A.4 Conclusion

While multiparametric EFRK-methods seem a useful tool at first sight, they come with many problems. Due to the strong dependence on the used frequencies and the complexity coherent with more stages, their use is best limited to problems for which the true frequencies are well known.

Referenties

- [1] J. C. Butcher, *The Numerical Analysis of Ordinary Differential Equations: Runge-Kutta and General Linear Methods*, Wiley-Interscience (1987).
- [2] M. Calvo, J.M. Franco, J.I. Montijano & L. Rández, *Sixth-order symmetric and symplectic exponentially fitted modified Runge–Kutta methods of Gauss type*, Computer Physics Communications 178 (2008), 732–744.
- [3] M. Calvo, J.M. Franco, J.I. Montijano & L. Rández, *Sixth-order symmetric and symplectic exponentially fitted Runge–Kutta methods of Gauss type*, Journal of Computational and Applied Mathematics 223 (2009), 387–398.
- [4] M. Calvo, J.M. Franco, J.I. Montijano & L. Rández, *Structure preservation of exponentially fitted Runge–Kutta methods*, Journal of Computational and Applied Mathematics 218 (2008), 421 – 434.
- [5] C. Chun & B. Neta, *Trigonometrically-fitted methods: a review*, Mathematics 7 (2019), 1197.
- [6] E. Hairer, C. Lubich & G. Wanner, *Geometric Numerical Integration: Structure Preserving Algorithms for Ordinary Differential Equations*, Springer-Verlag (2002).
- [7] L. Ixaru, *Numerical Methods for Differential Equations*, Reidel (1984).
- [8] J. Lambert, *Numerical Methods for Ordinary Differential Equations*, Wiley (1991).
- [9] A. París & L. Rández, *New embedded explicit pairs of exponentially fitted Runge-Kutta methods*, Journal of Computational and Applied Mathematics 234 (2010), 767–776.
- [10] B. Paternoster, *Present state-of-the-art in exponential fitting. A contribution dedicated to Liviu Ixaru on his 70th birthday*, Computer Physics Communications 183 (2012), 2499–2512.
- [11] J. Sanz-Serna & M. Calvo, *Numerical Hamiltonian Systems*, Chapman & Hall (1994).
- [12] M. Van Daele, *Gevorderde numerieke methoden*, Universiteit Gent (academiejaar 2018–2019).
- [13] M. Van Daele, *Wiskundige modellering*, Universiteit Gent (academiejaar 2017–2018), 51–66.
- [14] M. Van Daele, D. Hollevoet & G. Vanden Berghe, *Three-stage two-parameter symplectic, symmetric exponentially-fitted Runge-Kutta methods of Gauss type*, AIP Conference Proceedings (2010).
- [15] H. Van de Vyver, *A fourth-order symplectic exponentially fitted integrator*, Computer Physics Communications 174 (2006), 255–262.

- [16] G. Vanden Berghe & M. Van Daele, *Symplectic exponentially-fitted modified Runge-Kutta methods of the Gauss type: revisited*, in Recent Advances in Computational and Applied Mathematics (Theodore E. Simos, 2011), 289-306.
- [17] G. Vanden Berghe, M. Van Daele & H. Vande Vyver, *Exponential fitted Runge-Kutta methods of collocation type: fixed or variable knot points?*, Journal of Computational and Applied Mathematics 159 (2003), 217–239.

Lijst van programma's

Voor de resultaten in deze masterproef werd gebruik gemaakt van de volgende programma's:

Matlab

- Tweetrapsmethode
 - *RKs2.m*: implementatie van RKs2
 - *RKs2_implicit_theta.m*: contourplots van de vergelijking voor θ
 - *RKs2-coeff-graphs.m*: grafieken van de coëfficiënten
 - *RKs2-bepaal_thetas.m*: grafiek van θ
 - *RKs2-cons_theta.m*: plot rond het geval $\theta = \frac{\sqrt{6}}{3}$
 - *RKs2-specials.m*: plots rond de speciale gevallen van z_2
 - *NRtheta.m*: bepaling van θ
- Drietrapsmethode
 - *RKs3.m*: implementatie van RKs3
 - *RKs3_implicit_theta.m*: contourplots van de vergelijking voor θ
 - *RKs3-coeff-graphs.m*: grafieken van de coëfficiënten
 - *RKs3-bepaal_thetas.m*: grafiek van θ
 - *RKs2-cons_theta.m*: plot rond het geval $\theta = \frac{\sqrt{15}}{10}$
 - *RKs3-specials.m*: plots rond de speciale gevallen van z_2
 - *RKs3-instability_test.m*: illustratie van de instabiliteit voor bepaalde waarden van z_i
 - *NRtheta3.m*: bepaling van θ
 - *G.m*: hulpfunctie bij de bepaling van θ
- Viertrapsmethode
 - *RKs4.m*: implementatie van RKs4
 - *RKs4-coeff-graphs.m*: grafieken van de b_i -coëfficiënten
 - *RKs4-bepaal_thetas.m*: grafiek van θ
 - *NRtheta4.m*: bepaling van θ
- Numerieke problemen
 - *kepler_alphas.m*: keplerprobleem
 - *pert_kepler_alphas.m*: geperturbeerd keplerprobleem
 - *simple_alphas.m*: autonoom en lineair probleem

- *euler_alphas.m*: eulerprobleem
- *exact_sol.m*: referentie-oplossingen
- Overige
 - *eta.m*: implementatie van de functies van Ixaru
 - *ordercheck.m*: illustraties van de ordeverlaging door minimalisatie van de lte

Maple

- Tweetrapsmethode
 - *RKs2_error.mw*: bepaling en onderzoek van de lte
 - *RKs2_coeff.mw*: bepaling en onderzoek van de coëfficiënten
 - *Stability_RKs2.mw*: bepaling en onderzoek van de stabiliteitsveelterm
- Drietrapsmethode
 - *RKs3_error.mw*: bepaling en onderzoek van de lte
 - *RKs3_coeff.mw*: bepaling en onderzoek van de coëfficiënten
 - *Stability_RKs3.mw*: bepaling en onderzoek van de stabiliteitsveelterm
- Viertrapsmethode
 - *RKs4_coeff.mw*: bepaling en onderzoek van de coëfficiënten

