Universität Leipzig
Fakultät für Mathematik und Informatik
(Mathematisches Institut)

Diplomarbeit

Ordnungssterne und Ordnungspfeile

Leipzig, Dezember 2011

vorgelegt von

Gesa Ortgies
Studiengang Mathematik

Betreuender Hochschullehrer: Prof. Dr. Peter Kunkel
Mathematisches Institut, Abteilung Numerik

Inhaltsverzeichnis

1 Einleitung

> Ordnungssterne sind die Blumen, die das Unkraut aus dem schönen Garten der A-Stabilität vertreiben.
> Ordnungspfeile sind das Leiterspiel, das an dunklen Wintertagen, an denen man nicht den Garten besucht, gespielt wird.[1]

So ähnlich führte John Butcher auf einer Konferenz die beiden Themen zusammen, die dieser Arbeit ihren Titel geben.

Differentialgleichungen tauchen bei vielen physikalischen Berechnungen auf, und ihre Lösungen werden fast nur noch mit Computern numerisch berechnet. In den 1950ern wurde ein bestimmter Typ Differentialgleichungen identifiziert, die sogenannten „steifen“ Differentialgleichungen. Bei diesen Gleichungen führen viele explizite Verfahren zu starken Schwankungen.
Der 1963 von Germund Dahlquist definierte Begriff der A-Stabilität gibt ein wichtiges Kriterium für die Möglichkeiten eines numerischen Verfahrens, auch für steife Differentialgleichungen eine „gute“ Lösung zu finden.
Diese Begriffe und weitere Grundlagen werden in den Kapiteln 2 und 3 vorgestellt. Bei den Aussagen zur numerischen Thematik wurde sich an die grundlegenden Bücher „Numerical Methods for Ordinary Differential Equations“ von John Butcher,[4] und „Solving Ordinary Differential Equations I & II“ von Hairer, Wanner und Nørsett[10, 11] gehalten.

Bei Untersuchungen zur A-Stabilität sind Ernst Hairer, Gerhard Wanner und Syvert Nørsett 1978 auf Abbildungen gestoßen, die sie als „Ordnungssterne“ bezeichneten. Mit diesen können auf anschauliche und ästhetische Art Untersuchungen zur Ordnung und zur A-Stabilität eines Verfahrens gemacht werden.
Bereits in ihrer ersten Veröffentlichung 1978 sind ihnen mit Hilfe von Ordnungssternen einige wichtige Beweise gelungen.[9] In ihrem zweiten Band über das numerische Lösen gewöhnlicher Differentialgleichungen fassen Hairer und Wanner 1991 diese und einige weitere Ergebnisse noch einmal zusammen.[11]
Nørsett hat dagegen mit Arieh Iserles 1991 das Buch „Order Stars“ herausgebracht, das Ordnungssterne hauptsächlich als Teil der Approximationstheorie losgelöst von Differentialgleichungen betrachtet.[13]
In Kapitel 4 sind die in den genannten Werken zu findenden Sätze und Ergebnisse zu

[1] Frei übersetzt und zusammengefasst nach [6], für das Originalzitat siehe Anhang B, S. 66

Ordnungssternen bei Einschrittverfahren vorgestellt. Dabei lag das Hauptaugenmerk dieser Arbeit auf der Ausarbeitung der Beweise.

John Butcher hat in den letzten Jahren die Ordnungspfeile eingeführt. Er stellt sie mit kurzen Beweisen zu den Eigenschaften von Ordnungspfeilen in seinem oben bereits erwähnten Buch[4] vor, schrieb bereits einige weitere Veröffentlichungen zum Thema (z.B. [5]), in denen er Beweise mit Hilfe von Ordnungspfeilen präsentiert, und stellte Ordnungspfeile auf vielen Konferenzen[1,6] vor. Ordnungspfeile können ähnlich wie die Ordnungssterne genutzt werden zur Analyse von numerischen Verfahren.
Die Ergebnisse von Butcher zu Ordnungspfeilen bei Einschrittverfahren werden in dieser Arbeit in Kapitel 5 vorgestellt. Auch hier liegt wieder die Betonung auf den Beweisen der Sätze.

2 Grundlagen

Es werden in der gesamten Arbeit skalare Differentialgleichungen der Form

$$y'(x) = f(x, y(x)) \tag{2.1}$$

betrachtet, wobei $f : \mathbb{R} \times \mathbb{C} \mapsto \mathbb{C}$ eine in ihrer zweiten Komponente Lipschitz-stetige und in x stetige Funktion ist und ein Anfangswert

$$y(x_0) = y_0 \tag{2.2}$$

gegeben ist. Vektorielle Differentialgleichungen können auf diese Form zurückgeführt werden. Durch den Satz von Picard-Lindelöf ist mit diesen Voraussetzungen die Eindeutigkeit und Existenz einer lokalen Lösung gegeben. Diese lässt sich jedoch nicht immer analytisch finden, so dass in den meisten Fällen Näherungslösungen mit numerischen Verfahren berechnet werden. Im Folgenden werden einige wichtige Begriffe zu numerischen Methoden und die wichtigsten der in dieser Arbeit vorkommenden Verfahren kurz vorgestellt.

2.1 Numerische Verfahren

Ein **explizites Einschrittverfahren** zur Lösung von Differentialgleichungen verwendet zur Berechnung des Wertes von y an der Stelle $x_n = x_0 + nh$ nur den vorher bereits berechneten Wert y_{n-1}. Im Gegenzug dazu verwendet ein **implizites** Verfahren auch den Wert y_n, sodass die Lösung meist nicht direkt gefunden werden kann. Stattdessen muss z.B. zunächst das Newton-Verfahren verwendet werden. Im allgemeinen ist hier h die feste Schrittweite eines Verfahrens. Es kann manchmal auch sinnvoll sein, die Schrittweite zur Erreichung eines Kompromisses zwischen sowohl besserer Stabilität als auch kleinerem Rechenaufwand zu variieren. Darauf wird jedoch hier nicht näher eingegangen. Soweit nicht anders gesagt, ist der Wert von h in allen Schritten gleich groß.
Mehrschrittverfahren benutzen mehrere der vorher berechneten Werte von y, und bei **mehrstufigen** Verfahren handelt es sich um Verfahren, die den aktuellen Wert durch Unterteilung der Schrittweite in mehrere Stufen berechnen.

2.1.1 Lokaler Fehler und Ordnung

Als lokaler Fehler nach n Schritten wird bei einem numerischen Verfahren unter Annahme der Richtigkeit aller bisher ermittelten Werte die Differenz zwischen dem

tatsächlichen Wert $y(x_n)$ und dem numerisch ermittelten Wert y_n bezeichnet. Hat ein Einschrittverfahren die Iterationsvorschrift

$$y_n = y_{n-1} + h\Phi(x_{n-1}, y_{n-1}, x_n, y_n, h) \tag{2.3}$$

mit einer Funktion Φ, so sagt man, das Verfahren hat Ordnung p, wenn der lokale Fehler die Abschätzung

$$\|y_{n+1} - y(x_{n+1})\| \leq Ch^{p+1} \tag{2.4}$$

erfüllt. Dabei ist C eine Konstante.
Je höher die Ordnung, desto genauer kann das Ergebnis sein, allerdings bei deutlich erhöhtem Rechenaufwand. Im nächsten Kapitel wird außerdem gezeigt, dass sich aus hoher Ordnung nicht immer auf das Erreichen einer guten Lösung schließen lässt.

2.1.2 Das Euler-Verfahren

Das explizite Euler-Verfahren berechnet die Lösung der Differentialgleichung nach der folgenden Iterationsvorschrift:

$$y_{n+1} = y_n + hf(x_n, y_n) \tag{2.5}$$

Dabei bezeichnet h die Schrittweite. Es handelt sich um ein **Einschrittverfahren**.
Bei dem impliziten Euler-Verfahren ist dagegen die Rechenvorschrift gegeben durch

$$y_{n+1} = y_n + hf(x_{n+1}, y_{n+1}). \tag{2.6}$$

Die Euler-Verfahren nähern bei der Berechnung die Fläche unter der Funktion f durch ein Rechteck an, dabei verwendet das explizite Euler-Verfahren die linke und das implizite Euler-Verfahren die rechte Seite des Intervalles $[x_n, x_{n+1}]$ als Stützstelle.
Durch Taylorentwicklung von $y(x_{n+1}) = y(x_n + h)$ lässt sich zeigen, dass die Ordnung des expliziten Euler-Verfahrens $p = 1$ ist.
Beim impliziten Euler-Verfahren ergibt sich durch zusätzliche Taylorentwicklung von $f(x_{n+1}, y(x_{n+1}))$ die gleiche Ordnung.

2.1.3 Runge-Kutta-Verfahren

Die s-stufigen Runge-Kutta-Verfahren sind Einschrittverfahren. Sie sind durch die folgenden Rechenvorschriften definiert:

$$\begin{aligned} k_i &= f\left(x_n + hc_i, y_n + \sum_{j=1}^{s} a_{i,j}k_j\right), \quad i = 1, ..., s \\ y_{n+1} &= y_n + \sum_{j=1}^{s} b_j k_j \end{aligned} \tag{2.7}$$

Dabei sind die $a_{i,j}$, b_j und c_i reelle Koeffizienten. Sie werden häufig in einer Tabelle (sogenanntes Butcher-Schema) wie folgt dargestellt:

$$\begin{array}{c|cccc} c_1 & a_{1,1} & a_{1,2} & \dots & a_{1,s} \\ c_2 & a_{2,1} & a_{2,2} & \dots & a_{2,s} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s,1} & a_{s,2} & \dots & a_{s,s} \\ \hline & b_1 & b_2 & \dots & b_s \end{array} \tag{2.8}$$

Gilt $a_{i,j} = 0$ für $j \geq i$, so handelt es sich um ein explizites Verfahren. Die Koeffizienten in der Diagonale und darüber werden dann bei der Darstellung meist weggelassen. Um eine bestimmte Ordnung zu erhalten, müssen bestimmte Bedingungen für die Koeffizienten gelten, diese lassen sich durch Taylorentwicklungen bestimmen.
Ein bekanntes explizites Runge-Kutta-Verfahren ist Runge-Kutta-4 (Ordnung 4), das die folgende Koeffiziententabelle hat:

$$\begin{array}{c|cccc} 0 & & & & \\ \frac{1}{2} & \frac{1}{2} & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ 1 & 0 & 0 & 1 & \\ \hline & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{array} \tag{2.9}$$

Eine Spezialform der impliziten Runge-Kutta-Verfahren sind die DIRK-Verfahren, bei denen $a_{i,j} = 0$ für $j > i$ und mindestens ein $a_{i,i} \neq 0$. DIRK steht für „diagonal implizites Runge-Kutta“-Verfahren. Wenn auch noch alle Diagonalelemente gleich sind ($a_{i,i} = \gamma,\ i = 1, ..., s$) so wird von einem SDIRK(„singly diagonal implicit Runge-Kutta“)-Verfahren gesprochen. Für verschiedene Beispiele werden später SDIRK3 Methoden (Ordnung 3) betrachtet. Ihr Butcher-Schema sieht so aus:

$$\begin{array}{c|cc} \gamma & \gamma & \\ 1-\gamma & 1-2\gamma & \gamma \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array} \qquad \text{mit } \gamma = \frac{3 \pm \sqrt{3}}{6} \tag{2.10}$$

Die Euler-Verfahren sind **1-stufige** Runge-Kutta-Verfahren. Sie können auch mit Hilfe von Butcher-Schemen dargestellt werden. Links ist das Schema des expliziten, rechts des impliziten Euler-Verfahrens zu sehen.

$$\begin{array}{c|c} 0 & \\ \hline & 1 \end{array} \qquad \begin{array}{c|c} 1 & 1 \\ \hline & 1 \end{array} \tag{2.11}$$

2.1.4 Mehrschrittverfahren

In den Abschnitten 3.6, 4.6 und 5.4 wird kurz die Thematik des jeweiligen Kapitels im Fall von Mehrschritt- statt Einschrittverfahren umrissen. Dabei wird von der folgenden allgemeinen Gleichung für lineare Mehrschrittverfahren mit k Schritten ausgegangen:

$$\sum_{j=0}^{k} \alpha_{k-j} y_{n+1-j} = h \sum_{j=0}^{k} \beta_{k-j} f(x_{n+1-j}, y_{n+1-j}) \tag{2.12}$$

Mehrschrittverfahren haben mehrere Anfangswerte nötig. Die ersten k Werte einer solchen Methode können z.B. mit Einschrittverfahren berechnet werden.

2.2 Funktionentheoretische Grundlagen

Für die Beweise zu den Sätzen und Ergebnissen bei den Ordnungssternen (und Ordnungspfeilen) werden einige wichtige Definitionen und Sätze aus der Funktionentheorie benötigt, sie werden hier kurz wiedergegeben, sowie sie in den Büchern „Complex Analysis“ von Elias M. Stein und Rami Shakarchi, erschienen 2003,[14] und „Complex Variables and Applications“ von Rual V. Churchill, herausgegeben 1974,[7] dargestellt wurden.

Definition 2.1 *Eine komplexwertige Funktion f heißt **holomorph** auf dem Gebiet $\Omega \subset \mathbb{C}$, wenn sie in allen Punkten von Ω holomorph ist, d.h. wenn für alle $z_0 \in \Omega$ die **komplexe Ableitung***

$$\lim_{h \to 0} \frac{f(z_0 + h) - f(z_0)}{h}, \quad h \in \mathbb{C} \tag{2.13}$$

existiert. Eine Funktion, die holomorph ist, ist also auf ihrem gesamten Definitionsgebiet komplex differenzierbar.
*Eine Funktion heißt **analytisch**, wenn sie in jedem Punkt durch eine Potenzreihe darstellbar ist. Eine analytische Funktion kann auf ihrem Definitionsgebiet keine Pole besitzen.*

Für holomorphe Funktionen gilt, dass sie unendlich oft differenzierbar sind. Somit ist jede holomorphe Funktion analytisch und diese Begriffe sind bei komplexen Funktionen äquivalent. In dieser Arbeit werden häufig Funktionen betrachtet, die bis auf endlich viele Punkte (Polstellen) auf ganz $\mathbb{C}$ holomorph sind.
Die folgenden drei Sätze geben noch einige grundlegende Bestandteile der Funktionentheorie wieder. Sie wurden so nach dem Kapitel „Einführung in die Funktionentheorie“ aus dem Buch „Mathematik für Physiker“ von Helmut Kaul und Helmut Fischer, erschienen 2011, dargestellt.[8]

Satz 2.2 *Eine komplexe Funktion*

$$f : z = x + iy \mapsto u(x, y) + iv(x, y) \tag{2.14}$$

*ist genau dann holomorph, wenn u und v als reellwertige Funktionen auf dem Gebiet $\Omega \subset \mathbb{R}^2$ C^1-differenzierbar sind und die **Cauchy-Riemannschen Differentialgleichungen***

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}; \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \tag{2.15}$$

erfüllen. Es gilt dann:

$$\frac{d}{dz} f(x + iy) = \frac{\partial u}{\partial x}(x, y) + i\frac{\partial v}{\partial x}(x, y) = \frac{\partial v}{\partial y}(x, y) - i\frac{\partial u}{\partial y}(x, y) \tag{2.16}$$

Besonders wichtig sind in dieser Arbeit komplexe Kurvenintegrale. Eine C^1-Kurve oder Weg c wird durch die Parametrisierung

$$c : [t_0, t_1] \mapsto \mathbb{C}, \quad c(t) = x(t) + iy(t) \tag{2.17}$$

beschrieben. Ist $t_0 = t_1$ so wird c als geschlossene Kurve bezeichnet. Umparametrisierungen, die in der gleichen Orientierung verlaufen, werden ebenfalls mit c bezeichnet, entgegengesetzt verlaufende Umparametrisierungen mit $-c$.
Es wird später noch die Kettenregel für holomorphe Funktionen auf einer Kurve c benötigt. Dabei bezeichnet im Folgenden der Strich die Ableitung nach z und der Punkt die Ableitung nach t.

Satz 2.3 *Ist f holomorph in $\Omega \subset \mathbb{C}$ und $t \mapsto c(t)$ C^1-differenzierbar mit $c(t) \in \Omega$ so gilt:*

$$\frac{d}{dt} f(c(t)) = f'(c(t))\, \dot{c}(t) \tag{2.18}$$

Beweis: Es wird $f(x + iy) = u(x, y) + iv(x, y)$ und $c(t) = c_1(t) + ic_2(t)$ gesetzt. Mit der Kettenregel und den Cauchy-Riemannschen Differentialgleichungen ergibt sich (dabei wird abkürzend für $u(x(t), y(t)) = u$ geschrieben u.ä.):

$$\begin{aligned}
\frac{d}{dt} f(c(t)) &= \frac{d}{dt}(u(c_1(t), c_2(t)) + iv(c_1(t), c_2(t)) \\
&= \frac{\partial u}{\partial x} \cdot \dot{c}_1(t) + \frac{\partial u}{\partial y} \cdot \dot{c}_2(t) + i \left(\frac{\partial v}{\partial x} \cdot \dot{c}_1(t) + \frac{\partial v}{\partial y} \cdot \dot{c}_2(t) \right) \\
&= \left(\frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} \right) (\dot{c}_1(t) + i\dot{c}_2(t)) \\
&= f'(c(t))\dot{c}(t)
\end{aligned} \tag{2.19}$$

□

Das Integral über eine Kurve ist wie folgt definiert:

Definition 2.4 *Sei $f : \Omega \subset \mathbb{C} \mapsto \mathbb{C}$ stetig und c ein durch $c : [t_0, t_1] \mapsto \mathbb{C}$ gegebenes orientiertes C^1- Kurvenstück in Ω. Dann ist das komplexe Kurven- oder Wegintegral definiert durch:*

$$\int_c f(z)dz := \int_{t_0}^{t_1} f(c(t))\dot{c}(t)dt \tag{2.20}$$

Bei geschlossenen Kurven wird auch die Schreibweise

$$\oint_c f(z)dz \tag{2.21}$$

verwendet.

Als wichtiger Satz wird noch das sogenannte Argumentenprinzip benötigt:

Satz 2.5 *Sei f eine analytische Funktion auf und innerhalb einer einfachen geschlossenen Kurve c, bis auf höchstens endlich viele Pole innerhalb von c. Außerdem besitzt f maximal eine endliche Anzahl Nullstellen innerhalb des von c umschlossenen Gebiets und keine auf der Kurve selber. Dann gilt, wenn c positiv orientiert durchlaufen wird, folgende Gleichung:*

$$\frac{1}{2\pi i}\oint_c \frac{f'(z)}{f(z)}dz = N - P \tag{2.22}$$

Dabei werden die P Pole und N Nullstellen mit allen ihren Vielfachheiten gezählt.

Der Bruch $\frac{f'(z)}{f(z)}$ wird als logarithmische Ableitung bezeichnet, da er sich umschreiben lässt zu $\frac{f'(z)}{f(z)} = \frac{d}{dz}\log(f(z))$. Die Logarithmusfunktion ist im komplexen Raum nicht eindeutig definiert. Der Logarithmus von f lässt sich wiederum umschreiben zu:

$$\log(f(z)) = \log(|f(z)|) + i\arg(f(z)) \tag{2.23}$$

Dabei ist $\log(|f(z)|)$ der reelle Logarithmus einer positiven reellen Zahl, der eindeutig definiert ist, und $\arg(f(z))$ bezeichnet das Argument, also den Winkel von f, der nur bis auf Addition von ganzzahligen Vielfachen von 2π eindeutig ist. Die Ableitung, also der Quotient $\frac{f'}{f}$, ist aber eindeutig und $\oint \frac{f'(z)}{f(z)}dz$ kann interpretiert werden als die Gesamtänderung im Winkel von f, wenn z die Kurve c durchläuft.
Der Beweis des Satzes geht über den Residuensatz und ist in [7] zu finden.

Desweiteren wird das Maximumprinzip gebraucht:

Satz 2.6 *Sei f eine nichtkonstante holomorphe Funktion auf einem Gebiet Ω, dann kann f in Ω kein Maximum erreichen.*

Beweis: Für den Beweis wird der Satz der offenen Abbildungen verwendet. Er besagt, dass bei holomorphen Funktionen offene Mengen auf offene Mengen abgebildet werden. Hätte f also ein Maximum an der Stelle z_0 in Ω, dann gäbe es eine offene Umgebung $U(z_0) \subset \Omega$ die auf eine offene Menge abgebildet werden muss, da f holomorph ist. Also muss es ein $z \in U$ geben, sodass $f(z) > f(z_0)$, was zu einem Widerspruch führt, womit der Satz bewiesen ist. □

2.3 Verhalten einer rationalen Funktion im Unendlichen

In dieser Arbeit wird an einigen Stellen das Verhalten von rationalen Funktionen im Unendlichen betrachtet. In diesem Abschnitt wird dieses Verhalten hergeleitet.
Zunächst werden die in dieser Arbeit verwendeten Schreibweisen für Wachstum ins Unendliche in der komplexen Ebene erläutert. Der Ausdruck $z \to \infty$, $z \in \mathbb{C}$ wird verwendet, wenn $|z| \to \infty$ in $\mathbb{R}$. An manchen Stellen soll aber deutlich gemacht werden, dass für bestimmte Eigenschaften z in einer bestimmten Richtung ins Unendliche geht. Es wird $z \to -\infty$ geschrieben, wenn für $z = x + iy$, $x \in \mathbb{R}$, $y \in \mathbb{R}$ gilt, dass $x \to -\infty$ und $z \to +\infty$, wenn $x \to \infty$ und in beiden Fällen y beschränkt bleibt.

$R(z)$ sei eine rationale Funktion. Alle rationalen Funktionen sind holomorph auf $\mathbb{C}$ bis auf endlich viele Stellen. Also kann $R(z)$ in jedem Punkt $z_0 \in \mathbb{C}$ in eine Laurentreihe entwickelt werden. Dann ist

$$R(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} a_{-n} (z - z_0)^{-n} \tag{2.24}$$

in einem Kreisring um z_0. Der Term mit negativen Exponenten von z wird Hauptteil der Laurentreihe genannt und der Term mit positiven Exponenten wird Nebenteil genannt. Durch Polynomdivision lässt sich für jede rationale Funktion zeigen, dass der erste Term, der Nebenteil der Laurentreihe, endlich ist, also ein Polynom vom Grad $k \geq 0$ ist. Wird $z = \frac{1}{u}$ gesetzt und um $z_0 = 0$ entwickelt, so ergibt sich:

$$R\left(\frac{1}{u}\right) = \sum_{n=0}^{k} a_n \left(\frac{1}{u}\right)^n + \sum_{n=1}^{\infty} a_{-n} \left(\frac{1}{u}\right)^{-n} \tag{2.25}$$

Es handelt sich jetzt also für

$$R\left(\frac{1}{u}\right) = \tilde{R}(u) = \sum_{n=0}^{k} a_n u^{-n} + \sum_{n=1}^{\infty} a_{-n} u^n \tag{2.26}$$

um einen endlichen Hauptteil. Es gibt dann ein $\ell \in \mathbb{Z}$ und ein $K \in \mathbb{R}$, sodass

$$\tilde{R}(u) = K u^{-\ell} + \mathcal{O}\left(u^{-\ell+1}\right), \quad u \to 0 \tag{2.27}$$

und somit für $z = \frac{1}{u}$

$$R(z) = K z^{\ell} + \mathcal{O}\left(z^{\ell-1}\right), \quad z \to \infty \tag{2.28}$$

gilt. Die Koeffizienten in Gleichung (2.27) lassen sich wie folgt berechnen. Für die rationale Funktion wird $R(z) = \frac{P(z)}{Q(z)}$ mit $k := \deg(P)$ und $j := \deg(Q)$ gesetzt. Durch Ausklammern von zunächst z^{k-j} und dann $\frac{p_k}{q_j}$ aus dem Zähler kann $R(z)$ dann wie folgt umgeschrieben werden:

$$\begin{aligned} R(z) = \frac{P(z)}{Q(z)} &= \frac{p_k z^k + p_{k-1} z^{k-1} + ... + p_1 z + p_0}{q_j z^j + q_{j-1} z^{j-1} + ... + q_1 z + q_0} \\ &= z^{k-j} \frac{p_k z^j + p_{k-1} z^{j-1} + ... + p_1 z^{j-(k-1)} + p_0 z^{j-k}}{q_j z^j + q_{j-1} z^{j-1} + ... + q_1 z + q_0} \\ &= \frac{p_k}{q_j} z^{k-j} \frac{q_j z^j + \tilde{p}_{k-1} z^{j-1} + ... + \tilde{p}_1 z^{j-(k-1)} + \tilde{p}_0 z^{j-k}}{q_j z^j + q_{j-1} z^{j-1} + ... + q_1 z + q_0} \end{aligned} \tag{2.29}$$

Dabei ist $\tilde{p}_{k-m} = p_{k-m} \frac{q_j}{p_k}$ für $m = 1, ..., k$. Es wird nun der gesamte Nenner im Zähler einmal addiert und einmal subtrahiert. Dadurch kann wie folgt umgeschrieben werden:

$$R(z) = \frac{p_k}{q_j} z^{k-j} \left(1 + \frac{\hat{p}_{k-1} z^{j-1} + ... + \hat{p}_n z^n}{q_j z^j + q_{j-1} z^{j-1} + ... + q_1 z + q_0}\right) \tag{2.30}$$

Dabei ist $\hat{p}_{k-m} = \tilde{p}_{k-m} - q_{j-m}$ für $m = 1, 2, ..., n$ mit $n = \min\{(j-k), 0\}$. Ist $k = j$, so ist das klar definiert. Für $k > j$ ist noch als Ergänzung für $m > j$ notwendig, dass alle q mit negativem Index gleich Null gesetzt werden. Ebenso werden für $k < j$ und $m > k$ alle p mit negativem Index gleich Null gesetzt. Die Summe im Zähler geht dann vom höchsten Exponenten $j-1$ bis zum niedrigsten Exponenten n. Aus dem Bruch in der Klammer kann jetzt aus dem Nenner z ausgeklammert werden:

$$R(z) = \frac{p_k}{q_j}\, z^{k-j} \left(1 + \frac{1}{z}\, \frac{\hat{p}_{k-1} z^{j-1} + ... + \hat{p}_n z^n}{q_j z^{j-1} + q_{j-1} z^{j-2} + ... + q_1 + q_0 z^{-1}}\right) \tag{2.31}$$

Für $z \to \infty$ ist der Bruch in der Klammer rechts konstant. Also gilt

$$\begin{aligned} R(z) &= \frac{p_k}{q_j}\, z^{k-j} \left(1 + \frac{1}{z}\, \mathcal{O}(1)\right), \quad z \to \infty \\ &= \frac{p_k}{q_j}\, z^{k-j} + \mathcal{O}\left(z^{k-j-1}\right), \quad z \to \infty. \end{aligned} \tag{2.32}$$

Es ergibt sich also

$$K = \frac{p_k}{q_j} \quad \text{und} \quad \ell = k - j. \tag{2.33}$$

3 Stabilität und steife Differentialgleichungen

3.1 Grundbegriffe und A-Stabilität

Bei einigen Differentialgleichungen tritt bei Anwendung von expliziten Verfahren ein Phänomen auf, dass sich häufig durch starke Oszillationen der numerischen Lösung um die exakte Lösung äußert. Solche Differentialgleichungen werden als **steif** bezeichnet. Es gibt keine sehr präzisen Definitionen für diesen Begriff. In [3] wird ausführlich über verschiedene bis dahin veröffentlichte Definitionen diskutiert. Die folgende Definition von J.D. Lambert charakterisiert den Effekt, den steife Differentialgleichungen auf manche Verfahren haben:

„If a numerical method with a finite region of absolute stability, applied to a system with any initial condition, is forced to use in a certain interval of integration a step length which is excessively small in relation to the smoothness of the exact solution in that interval, then the system is said to be stiff in that interval“[2].

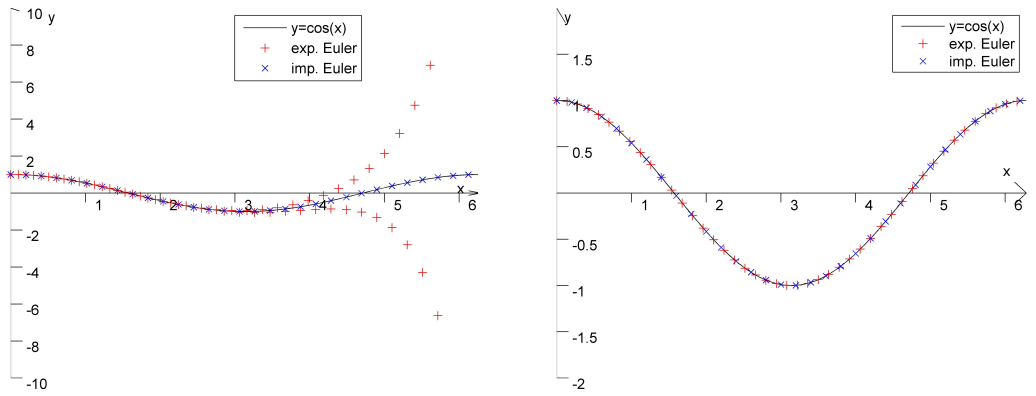

Abbildung 3.1: Explizites vs. implizites Euler-Verfahren an Beispielfunktion (3.1). Links $h = 1/49$, rechts $h = 1/50$

Dies soll durch ein Beispiel verdeutlicht werden. Es wird das Anfangswertproblem

$$y'(x) = -100(y(x) - \cos(x)) - \sin(x), \quad y(0) = 1 \tag{3.1}$$

[2]aus: J.D. Lambert: „Numerical Methods for Ordinary Differential Equations“, *John Wiley, 1991*, zitiert nach [3]

betrachtet. Die exakte Lösung hiervon ist $y(x) = \cos(x)$.
In Tabelle 3.1 wird der Wert des lokalen Fehlers $|y(x) - \tilde{y}(x)|$ an bestimmten Stellen x für zwei verschiedene Werte von h dargestellt. Dabei bezeichne $\tilde{y}$ die durch das jeweilige Verfahren berechneten Werte der Funktion.

	$h = \frac{1}{49}$		$h = \frac{1}{50}$	
x	exp. Euler	imp. Euler	exp. Euler	imp. Euler
1	0,000780	0,000057	0,000046	0,000055
2	0,005188	0,000041	0,000141	0,000040
3	0,036440	0,000101	0,000199	0,000099
4	0,259552	0,000068	0,000166	0,000067
5	1,842678	0,000027	0,000072	0,000027
6	13,084896	0,000097	0,000004	0,000096

Tabelle 3.1: $|y(x) - \tilde{y}(x)|$ berechnet für die Beispielfunktion aus Gleichung (3.1)

Die Abbildung 3.1 stellt die berechneten Funktionswerte im Vergleich zur exakten Lösung noch einmal graphisch dar. Im linken Bild sind Lösungen für $h = \frac{1}{49}$ zu sehen, die Werte für das implizite Euler-Verfahren stimmen näherungsweise mit denen der exakten Lösung überein, während die Lösungen des expliziten Euler-Verfahrens immer mehr schwanken. Im rechten Bild für $h = \frac{1}{50}$ stimmt auch die Lösung des expliziten Euler-Verfahrens gut mit der exakten Lösung überein. Der lokale Fehler muss nach Formel (2.4), da die Euler-Verfahren die Ordnung 1 haben, kleiner sein als Ch^2 für eine Konstante C. Dies ist beim expliziten Euler-Verfahren für $h = \frac{1}{49}$ nicht gegeben, aber für $h = \frac{1}{50}$ bleibt der Fehler beschränkt.
Die Lösung für $h = \frac{1}{49}$ mit dem expliziten Euler-Verfahren ist instabil, und das obwohl sich der Wert von h nur gering von dem für eine stabile Lösung unterscheidet. Wie es zu dieser Unterscheidung kommt, wird nun in diesem Kapitel erklärt.

Ob ein numerisches Verfahren für die Lösung einer steifen Differentialgleichung geeignet ist, kann untersucht werden an der sogenannten *Dahlquistschen Testgleichung*:

$$y'(x) = \lambda y(x), \qquad y(0) = 1 \tag{3.2}$$

Die exakte Lösung hierfür ist $y(x) = e^{\lambda x}$. Für $\mathrm{Re}(\lambda) \leq 0$ ist $y(x)$ für $x \to \infty$ beschränkt. Dieses Verhalten sollte auch die numerische Lösung aufweisen. Im $(n+1)$-ten Schritt hängt y_{n+1} für ein Runge-Kutta-Verfahren mit der vorherigen Lösung für y_n über $z = h\lambda$ wie folgt zusammen: Es gilt

$$y_{n+1} = R(z)y_n \tag{3.3}$$

und somit $y_{n+1} = R(z)^n y_0$. Dies ist beschränkt, falls $|R(z)| \leq 1$. Daraus leiten sich die folgenden zwei Definitionen ab:

Definition 3.1 *$R(z)$ wird bezeichnet als Stabilitätsfunktion eines Verfahrens.*

Definition 3.2 *Die Menge*

$$S = \{z \in \mathbb{C} : |R(z)| \leq 1\} \tag{3.4}$$

wird als Stabilitätsbereich eines Verfahrens bezeichnet.

Für ein Verfahren, dessen Stabilitätsfunktion für alle λ mit $\mathrm{Re}(\lambda) \leq 0$ durch 1 beschränkt ist, wurde die Bezeichnung A-Stabilität eingeführt.

Definition 3.3 *(Dahlquist, 1963) Ein Verfahren, bei dem der Stabilitätsbereich die linke Halbebene $\mathbb{C}^-$ beinhaltet, heißt A-stabil.*

In dieser Arbeit wird häufig von einer A-stabilen (Stabilitäts-)Funktion gesprochen. Gemeint ist damit, dass das zugehörige Verfahren A-stabil ist.

3.2 Beispiele

Bei dem expliziten Euler-Verfahren gilt die Iterationsvorschrift aus Gleichung (2.5). Diese lässt sich für die Testgleichung (3.2) umschreiben zu

$$y_{n+1}(x) = (1 + h\lambda)y_n(x). \tag{3.5}$$

Die Stabilitätsfunktion lautet also $R(z) = 1 + z$ mit $z = h\lambda$ und der Stabilitätsbereich ist die Kreisscheibe mit Radius 1 um den Punkt $(-1, 0)$, beschrieben durch

$$\{z \in \mathbb{C} : |1 + z| < 1\}. \tag{3.6}$$

Für negative Realteile von λ ergibt sich also $h < \frac{2}{|\lambda|}$, was für große Werte von $|\lambda|$ eine sehr kleine nötige Schrittweite bedeutet. Das Anfangswertproblem am Anfang dieses Kapitels verhält sich näherungsweise wie $y'(x) = -100y(x)$. Dadurch erklärt sich der gewählte Grenzwert $h = \frac{1}{50} = \frac{2}{|-100|}$.
Bei dem impliziten Euler-Verfahren ergibt sich dagegen durch Einsetzen der Testgleichung in die Iterationsvorschrift in Gleichung (2.6) und Umformen die folgende Beziehung:

$$y_{n+1}(x) = \frac{1}{1 - h\lambda}y(x) \tag{3.7}$$

Der Stabilitätsbereich

$$\left\{z \in \mathbb{C} : \left|\frac{1}{1 - z}\right| < 1\right\} \tag{3.8}$$

enthält somit die komplette linke Halbebene, wodurch h für $\mathrm{Re}(\lambda) < 0$ zunächst beliebig gewählt werden kann. Das implizite Euler-Verfahren ist also A-stabil.
Abbildung 3.2 zeigt farbig die Stabilitätsbereiche der beiden Euler-Verfahren. Dabei geht die Farbskala von dunkelblau für Absolutwerte nahe 1 zu hellblau für kleine Absolutwerte nahe 0.
Bei den Runge-Kutta-Verfahren ist die Stabilitätsfunktion immer eine rationale Funktion, wobei es sich bei den expliziten Verfahren mit s Stufen um ein Polynom vom

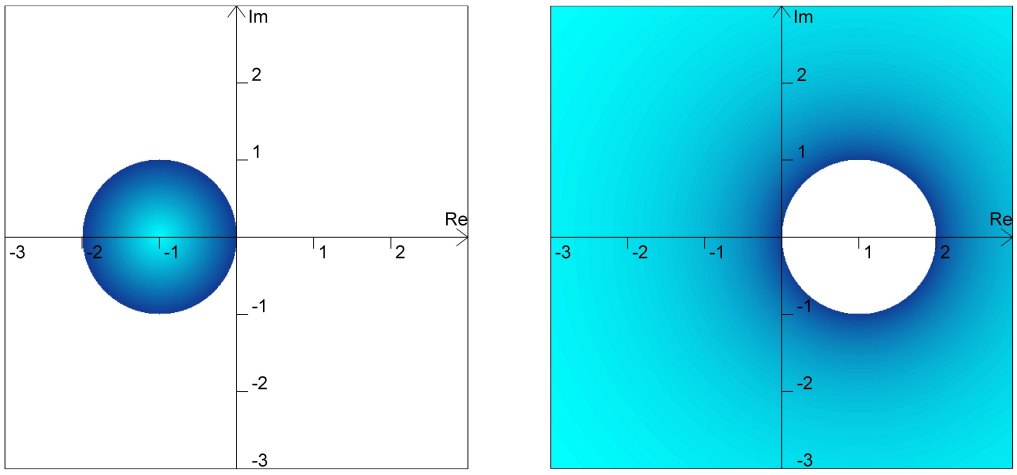

Abbildung 3.2: Stabilitätsbereich explizites (links) und implizites Euler-Verfahren

Grad $\leq s$ handelt. Wenn die Ordnung des Verfahrens gleich der Anzahl der Stufen ist ($p = s$), dann ist die Stabilitätsfunktion eines expliziten Runge-Kutta-Verfahrens gegeben durch:

$$R(z) = 1 + z + \frac{z^2}{2!} + ... + \frac{z^s}{s!} \tag{3.9}$$

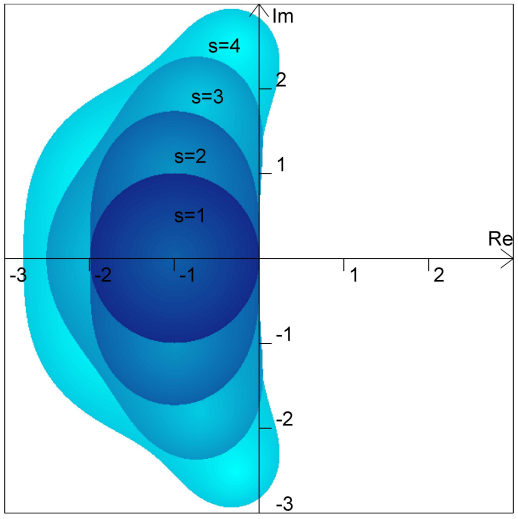

Abbildung 3.3: Stabilitätsbereiche expliziter Runge-Kutta-Verfahren

Für $s = 1, 2, 3$ und 4 sind die Stabilitätsbereiche in Abbildung 3.3 gezeigt. Je höher die Ordnung, desto größer ist auch der Stabilitätsbereich der Runge-Kutta-Verfahren.

Allgemein ist bei Runge-Kutta-Verfahren die Stabilitätsfunktion gegeben durch:

$$R(z) = \frac{\det(I - zA + z\mathbb{1}b^T)}{\det(I - zA)} \tag{3.10}$$

Dabei ist I die Identitätsmatrix, $\mathbb{1}$ bezeichnet den Vektor aus Einsen, und A und b sind die Koeffizienten des Verfahrens (siehe Schema in (2.8)). Nicht alle impliziten Verfahren sind A-stabil.
Bei den SDIRK3-Verfahren ist die Stabilitätsfunktion demnach

$$R(z) = \frac{1 + z(1 - 2\gamma) + z^2(\frac{1}{2} - 2\gamma + \gamma^2)}{(1 - \gamma z)^2}. \tag{3.11}$$

Für positives Vorzeichen vor der Wurzel in γ (siehe Gleichung (2.10)) ist A-Stablität gegeben, für negatives aber nicht. Beides sind implizite Verfahren der Ordnung 3.
In der Abbildung 3.4 sind die Stabilitätsgebiete der verwendeten SDIRK3-Verfahren dargestellt, links für $\gamma = \frac{3+\sqrt{3}}{6}$ und rechts für $\gamma = \frac{3-\sqrt{3}}{6}$.

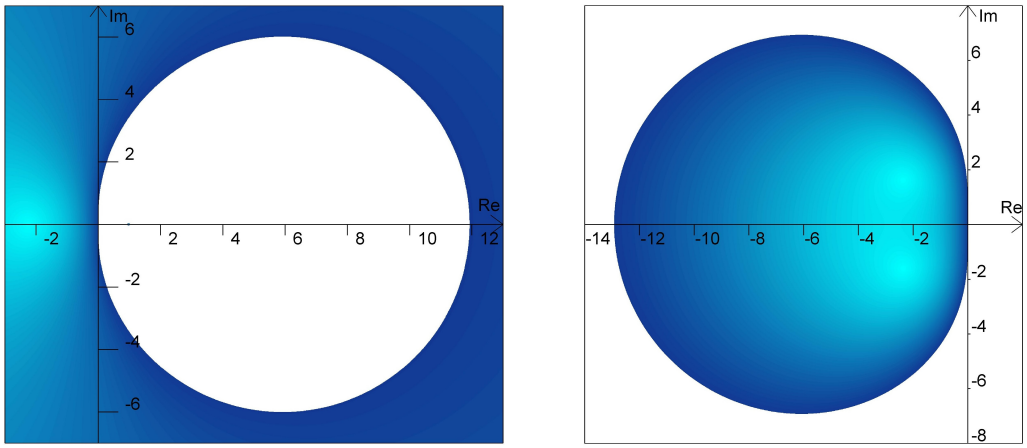

Abbildung 3.4: Stabilitätsbereich SDIRK3-Verfahren

3.3 Padé-Approximationen

Die Stabilitätsfunktionen aller behandelten Funktionen sind mehr oder weniger gute Approximationen an die Exponentialfunktion.

Definition 3.4 *$R(z)$ ist eine Approximation an e^z der Ordnung p, wenn es eine Konstante C gibt, so dass gilt:*

$$e^z - R(z) = Cz^{p+1} + \mathcal{O}(z^{p+2}), \quad z \to 0 \tag{3.12}$$

Im Allgemeinen wird von $C \neq 0$ ausgegangen und die höchste Ordnung von R angegeben, d.h. für $C = 0$ wird die Ordnung $p + 1$ verwendet. Wird $R(z) = \frac{P(z)}{Q(z)}$ gesetzt und ist der Grad von P und Q jeweils $\leq s$, dann folgt für eine Approximation der Ordnung $\geq s$ hieraus

$$e^z Q(z) = P(z) + C_1 z^{s+1} + C_2 z^{s+2} + ... \tag{3.13}$$

mit Konstanten C_1, C_2 usw. Ist $Q(z)$ bekannt, so sind $P(z)$ und auch die Konstanten eindeutig bestimmt. Ist $Q(z)$ gegeben durch

$$Q(z) = q_0 + q_1 z + q_2 z^2 ... + q_s z^s \tag{3.14}$$

so gilt, da $e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + ...$:

$$P(z) = q_0 + \left(\frac{q_0}{1!} + \frac{q_1}{0!}\right) z + \left(\frac{q_0}{2!} + \frac{q_1}{1!} + \frac{q_2}{0!}\right) z^2 + ... + \left(\frac{q_0}{s!} + \frac{q_1}{(s-1)!} + ... + \frac{q_s}{0!}\right) z^s \tag{3.15}$$

Über diese Gleichungen kann mit einigen weiteren Hilfsschritten, die ausführlich in [11] dargestellt sind, die Padé-Approximationen bestimmt werden.
Padé-Approximationen sind rationale Funktionen von höchstmöglicher Ordnung $p = j + k$ der Approximation, wobei j der Grad des Nenners und k der Grad des Zählers der Funktion ist. Die Padé(k, j)-Approximation an e^z ist gegeben durch

$$R_{kj}(z) = \frac{P_{kj}(z)}{Q_{kj}(z)} \tag{3.16}$$

mit

$$P_{kj}(z) = 1 + \frac{k}{j+k} z + \frac{k(k-1)}{(j+k)(j+k-1)} \cdot \frac{z^2}{2!} + ... + \frac{k(k-1)...1}{(j+k)...(j+1)} \cdot \frac{z^k}{k!} \tag{3.17}$$

$$Q_{kj}(z) = 1 - \frac{j}{k+j} z + \frac{j(j-1)}{(k+j)(k+j-1)} \cdot \frac{z^2}{2!} + ... + (-1)^j \frac{j(j-1)...1}{(k+j)...(k+1)} \cdot \frac{z^j}{j!}. \tag{3.18}$$

Nullstellen von P_{kj} sind Nullstellen der Padé-Approximation und Nullstellen von Q_{kj} sind Pole der Padé-Approximation.
In der Tabelle in (3.19) sind einige Beispiele von Padé-Approximationen gegeben. Es fällt auf, dass viele der behandelten Verfahren eine Padé-Approximation als Stabilitätsfunktion haben – sie sind also für ihre Ordnung bereits optimale Approximationen. Es ist außerdem hier die Ordnung der Approximation gleichwertig mit der Ordnung des Lösungsverfahrens.

$j \backslash k$	0	1	2	3
0	1	$1 + z$	$1 + z + \frac{1}{2}z^2$	$1 + z + \frac{1}{2}z^2 + \frac{1}{6}z^3$
1	$\frac{1}{1-z}$	$\frac{1+\frac{1}{2}z}{1-\frac{1}{2}z}$	$\frac{1+\frac{2}{3}z+\frac{1}{6}z^2}{1-\frac{1}{3}z}$	$\frac{1+\frac{3}{4}z+\frac{1}{4}z^2+\frac{1}{24}z^4}{1-\frac{1}{4}z}$
2	$\frac{1}{1-z+\frac{1}{2}z^2}$	$\frac{1+\frac{1}{3}z}{1-z+\frac{1}{2}z^2}$	$\frac{1+\frac{1}{2}z+\frac{1}{12}z^2}{1-\frac{1}{2}z+\frac{1}{12}z^2}$	$\frac{1+\frac{3}{5}z+\frac{3}{20}z^2+\frac{1}{60}z^3}{1-\frac{2}{5}z+\frac{1}{20}z^2}$
3	$\frac{1}{1-z+\frac{1}{2}z^2-\frac{1}{6}z^3}$	$\frac{1+\frac{1}{4}z}{1-\frac{3}{4}z+\frac{1}{4}z^2-\frac{1}{24}z^3}$	$\frac{1+\frac{2}{5}z+\frac{1}{20}z^2}{1-\frac{3}{5}z+\frac{3}{20}z^2-\frac{1}{60}z^3}$	$\frac{1+\frac{1}{2}z+\frac{1}{10}z^2+\frac{1}{120}z^3}{1-\frac{1}{2}z+\frac{1}{10}z^2-\frac{1}{120}z^3}$

(3.19)

So ist z.B. die Stabilitätsfunktion der Euler-Verfahren jeweils im expliziten Fall durch Padé(1,0) und im impliziten Fall durch Padé(0,1) gegeben und die Stabilitätsfunktion von Runge-Kutta-4 ist Padé(4,0).

3.4 Das E-Polynom

Das E-Polynom ist definiert für rationale Stabilitätsfunktionen und wurde von Nørsett eingeführt zur Analyse der Stabilität eines Verfahrens.

Definition 3.5 *Für eine rationale Funktion $R(z) = \frac{P(z)}{Q(z)}$ ist das E-Polynom definiert als:*

$$E(y) = |Q(iy)|^2 - |P(iy)|^2 \tag{3.20}$$

Es wird später eine spezielle Eigenschaft des E-Polynoms gebraucht, die hier beschrieben wird.

Satz 3.6 *$E(y)$, wie oben definiert, ist ein gerades Polynom. Es hat den Grad*

$$\deg(E) \leq 2 \max\{\deg P, \deg Q\}.$$

Wenn $R(z)$ eine Approximation der Ordnung p ist, dann gilt:

$$E(y) = \mathcal{O}(y^{p+1}), \quad y \to 0 \tag{3.21}$$

Beweis: Es folgt direkt aus der Definition, dass E gerade ist. Nach der Definition 3.4 gilt:

$$\begin{aligned} & e^z - R(z) = \mathcal{O}(z^{p+1}), \quad z \to 0 \\ \Leftrightarrow \quad & |\, e^z - R(z)\,| \leq M \left|z^{p+1}\right| \end{aligned} \tag{3.22}$$

Dabei ist M eine Konstante. Mit der umgekehrten Dreiecksungleichung

$$||a| - |b|| \leq |a - b| \tag{3.23}$$

ergibt sich:

$$\begin{aligned} \Rightarrow \quad & |\,|e^z| - |R(z)|\,| \leq M \left|z^{p+1}\right| \\ \Leftrightarrow \quad & |e^z| - |R(z)| = \mathcal{O}(z^{p+1}), \quad z \to 0 \\ \Leftrightarrow \quad & |e^z| - \left|\frac{P(z)}{Q(z)}\right| = \mathcal{O}(z^{p+1}), \quad z \to 0 \end{aligned} \tag{3.24}$$

Wird nun $z = iy$, $y \in \mathbb{R}$ eingesetzt so gilt, da $|e^{iy}| = 1$ und da Q eine ganzrationale Funktion ist und daher $|Q(iy)| = \mathcal{O}(1),\ y \to 0$:

$$|Q(iy)| - |P(iy)| = \mathcal{O}(y^{p+1}), \quad y \to 0 \tag{3.25}$$

Es gilt

$$\begin{aligned} E(y) &= |Q(iy)|^2 - |P(iy)|^2 \\ &= (|Q(iy)| + |P(iy)|)(|Q(iy)| - |P(iy)|) \\ &= \mathcal{O}(1)\mathcal{O}(y^{p+1}), \quad y \to 0 \end{aligned} \tag{3.26}$$

und damit folgt die Behauptung. □

3.5 Weitere Stabilitätsbegriffe

A-Stabilität ist nicht immer ausreichend. Es gibt Differentialgleichungen, die so steif sind, dass auch einige A-stabile Verfahren starke Schwankungen in den Werten aufweisen. Daher ist es manchmal wünschenswert, dass der Wert von $|R(z)|$ für $z \to -\infty$ deutlich kleiner als 1 ist. Hieraus wurde der Begriff der L-Stabilität hergeleitet.

Definition 3.7 *Ein Verfahren heißt L-stabil, wenn es A-stabil ist und zusätzlich Folgendes gilt:*

$$\lim_{z \to \infty} R(z) = 0 \tag{3.27}$$

Gleichzeitig ist aber auch für viele Gleichungen wirkliche A-Stabilität gar nicht nötig, und es gibt auch die abgeschwächte Version der $A(\alpha)$-Stabilität, die meist ausreichend ist:

Definition 3.8 *Ein Verfahren wird $A(\alpha)$-stabil genannt, wenn für einen Winkel $\alpha \in [0, \frac{\pi}{2}]$ alle $z = re^{i\theta}$, für die die folgenden zwei Bedingungen gelten, im Stabilitätsbereich des Verfahrens enthalten sind.*

(i) $\theta \in \left[\frac{\pi}{2}, \frac{3\pi}{2}\right]$

(ii) $|\theta - \pi| \leq \alpha$

Als Beispiel zeigt Bild 3.5 den Stabilitätsbereich der Padé(0,4)-Approximation, die $A(\alpha)$-stabil ist mit $\alpha \sim 83,7°$.

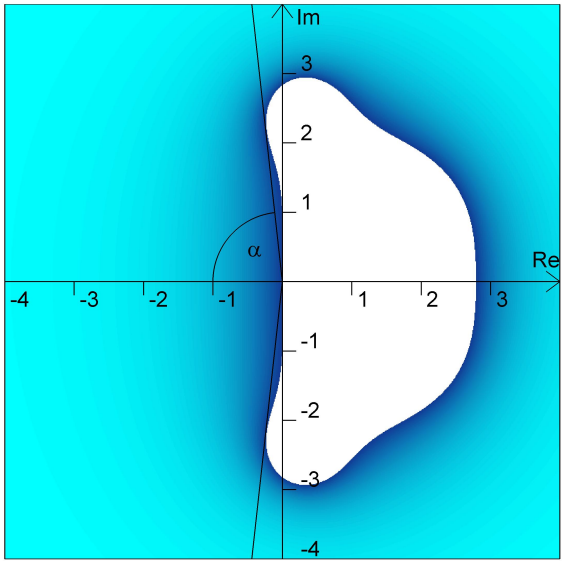

Abbildung 3.5: Beispiel für $A(\alpha)$-Stabilität

3.6 Stabilität bei Mehrschrittverfahren

Bei Mehrschrittverfahren ist die Untersuchung der Stabilität eines Verfahrens weitaus komplizierter. Daher wird hier nur kurz als Ausblick auf die Bezeichnungen eingegangen. Wird die Dahlquist'sche Testgleichung auf Gleichung (2.12) angewendet, so ergibt sich nach Umformen:

$$(\alpha_k - h\lambda\beta_k)y_{n+1} + (\alpha_{k-1} - h\lambda\beta_{k-1})y_n + ... + (\alpha_0 - h\lambda\beta_0)y_{n+1-k} = 0 \tag{3.28}$$

Nach der Methode von Lagrange wird $y_j = \zeta^{j+k-n-1}$ gesetzt. Wie bei den Einschrittverfahren wird $z = h\lambda$ gesetzt. Es ergibt sich dann das charakteristische Polynom:

$$(\alpha_k - z\beta_k)\zeta^k + (\alpha_{k-1} - z\beta_{k-1})\zeta^{k-1} + ... + (\alpha_0 - z\beta_0)\zeta^0 = 0 \tag{3.29}$$

Nun wird der Stabilitätsbereich eines Mehrschrittverfahrens wie folgt definiert:

Definition 3.9 *Als Stabilitätsbereich eines Mehrschrittverfahrens oder als Bereich absoluter Stabilität wird die Menge*

$$S = \left\{ z \in \mathbb{C} : \begin{array}{l} \textit{Für alle Nullstellen } \zeta_j(z) \textit{ gilt } |\zeta_j(z)| \leq 1, \\ \textit{mehrfache Nullstellen erfüllen } |\zeta_j(z)| < 1 \end{array} \right\} \tag{3.30}$$

bezeichnet. Das Verfahren ist A-stabil, wenn $S \supset \mathbb{C}^-$.

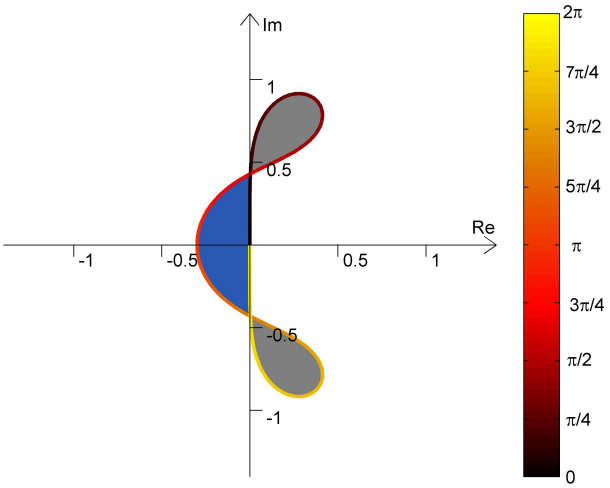

Abbildung 3.6: Stabilitätsbereich auf z-Ebene, Adams-Bashforth Verfahren der Ordnung 4

Dabei ist j ein Index, der die verschiedenen Nullstellen angibt.
Werden die Nullstellen $\zeta_j(z)$ auf der ζ-Ebene betrachtet, so müssen sie für A-Stabilität alle innerhalb des Einheitskreises $\zeta = e^{i\theta}$ liegen. Dieser Einheitskreis wird als positiv orientierte Kurve von $\theta_0 = 0$ bis $\theta_1 = 2\pi$ betrachtet und auf die z-Ebene abgebildet, indem die obige Formel (3.29) nach z umgeformt wird.

Da das Innere des Einheitskreises links von der gegebenen Kurve liegt, gilt das auch für die z-Ebene. Nur die Bereiche sind Teil des Stabilitätsgebietes, die von positiv durchlaufenen Kurvenstücken umrahmt werden.
Als Beispiel zeigt Abbildung 3.6 das Stabilitätsgebiet der Adams-Bashforth-Verfahrens der Ordnung 4 (ein explizites Mehrschrittverfahren), dessen Iterationsvorschrift

$$y_{n+1} = y_n + h\left(\frac{55}{24}f(x_n, y_n) - \frac{59}{24}f(x_{n-1}, y_{n-1}) + \frac{37}{24}f(x_{n-2}, y_{n-2}) - \frac{9}{24}f(x_{n-3}, y_{n-3})\right) \tag{3.31}$$

das charakteristische Polynom

$$\zeta^4 - \left(1 + \frac{55}{24}z\right)\zeta^3 + \frac{59}{24}z\zeta^2 - \frac{37}{24}z\zeta + \frac{9}{24}z = 0 \tag{3.32}$$

hat, was sich umformen lässt zu

$$z = \frac{\zeta^4 - \zeta^3}{\frac{55}{24}\zeta^3 - \frac{59}{24}\zeta^2 + \frac{37}{24}\zeta - \frac{9}{24}}. \tag{3.33}$$

Für die Abbildung wurde in die Formel für z die Gleichung $\zeta = e^{i\theta}$ eingesetzt. Diese Kurve, die als Wurzelortskurve bezeichnet wird, ist im Bild mit aufsteigendem Wert von θ farbig dargestellt, dadurch ist erkennbar, welcher Bereich positiv umlaufen wird. Nur der blaue Bereich gehört zum Stabilitätsgebiet des Verfahrens.

4 Ordnungssterne

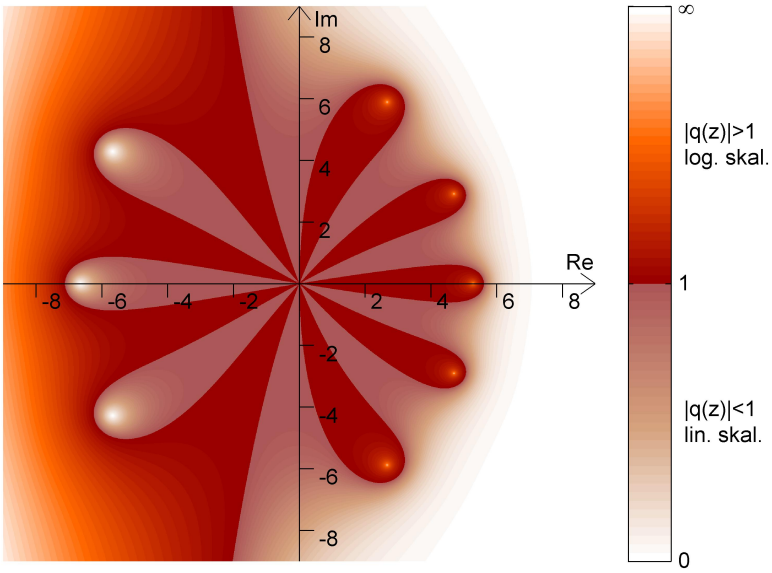

Abbildung 4.1: Ordnungsstern Padé(3,5)

4.1 Eigenschaften von Ordnungssternen

Bei der Betrachtung der Stabilititätsgebiete von Padé-Approximationen haben E. Hairer, G. Wanner und S.P. Nørsett 1978 in [9] den Begriff Ordnungsstern eingeführt. Ordnungssterne verdeutlichen auf eine anschauliche Weise die Ordnung und Stabilität eines Verfahrens.
Bisher wurde der Betrag von der Stabilitätsfunktion $R(z)$ mit 1 verglichen, um zu sehen, ob die numerische Lösung wie die Exponentialfunktion beschränkt ist, in den Ordnungssternen wird aber nun die Stabilitätsfunktion direkt mit der Exponentialfunktion verglichen:

Definition 4.1 *Die Menge*

$$A = \left\{ z \in \mathbb{C} : \left| \frac{R(z)}{e^z} \right| > 1 \right\} \tag{4.1}$$

wird als Ordnungsstern bezeichnet.

Dabei fällt auf, dass die Stabilitätsfunktion $R(z)$ und die Ordnungssternfunktion

$$q(z) := \frac{R(z)}{e^z} \tag{4.2}$$

die gleichen Nullstellen und Pole haben, da die Division mit der Exponentialfunktion keine Auswirkung hat auf diese Punkte.
In der Abbildung 4.1 ist bereits ein Beispiel eines Ordnungssternes der Padé(3,5)-Approximation aufgezeigt. Es wird immer von reellen Koeffizienten der Stabilitätsfunktion ausgegangen, dadurch ist die Symmetrie des Ordnungssterns zur x-Achse gegeben.
Es gibt alternative Definitionen der Ordnungssterne. In [13] definieren Iserles und Nørsett den Ordnungsstern 1. Art als Tupel aus den drei Mengen

$$\begin{aligned} \mathcal{A}_+ &= \{z \in \mathbb{C} : |q(z)| > 1\} \\ \mathcal{A}_0 &= \{z \in \mathbb{C} : |q(z)| = 1\} \\ \mathcal{A}_- &= \{z \in \mathbb{C} : |q(z)| < 1\}. \end{aligned} \tag{4.3}$$

Dies ändert jedoch nichts an den Eigenschaften und in dieser Arbeit wird immer die Definition 4.1 von A genutzt.
Die folgende Definition listet die in dieser Arbeit verwendeten Bezeichnungen für verschiedene Teile des Ordnungssterns auf.

Definition 4.2 *Die Komplementärmenge $\mathbb{C} \setminus A$ wird als **dualer Stern** bezeichnet.*
*Als **Sektor** werden die annähernd Kreissektor-förmigen Abschnitte von A und $\mathbb{C} \setminus A$ bezeichnet, die entstehen, wenn ein Kreis mit ausreichend kleinem Radius um den Ursprung gelegt wird (vergleiche dazu auch Satz 4.4).*
*Für die zu A gehörigen Zusammenhangskomponenten haben Hairer, Wanner und Nørsett die Bezeichnung **Finger** geprägt, die zum dualen Stern gehörigen Zusammenhangskomponenten bezeichnen sie als **duale Finger**. Wenn m der um den Ursprung herum separaten Sektoren sich weiter außen zu einem Finger zusammen finden (siehe Beispiel Bild 4.4 auf Seite 31) so wird dies als Finger der **Vielfachheit** m bezeichnet, analog gibt es auch duale Finger der Vielfachheit m (vergleiche Satz 4.7).*
*Als **Randlinien** von A werden die Elemente von ∂A bezeichnet, also die Linien, auf denen $|q(z)| = 1$. Über die Randlinien werden in den Sätzen 4.5 und 4.6 Aussagen getroffen.*

Es folgen einige Sätze, die die Eigenschaften von Ordnungssternen genauer aufzeigen. Der erste hiervon gibt ein Kriterium für die Erkennbarkeit von A-Stabilität anhand von Ordnungssternen.

Satz 4.3 *$R(z)$ ist A-stabil genau dann, wenn Folgendes gilt:*

(i) $A \cap i\mathbb{R} = \emptyset$

(ii) Alle Pole von $R(z)$ liegen in der positiven Halbebene $\mathbb{C}^+$.

Beweis: Es wird zunächst gezeigt, dass aus (i) und (ii) A-Stabilität folgt. Aus der ersten Bedingung folgt für alle $y \in \mathbb{R}$, dass $|q(iy)| = |R(iy)e^{-iy}| \leq 1$, woraus wiederum $|R(iy)| \leq 1$ folgt, da $|e^{-iy}| = 1$. Die imaginäre Achse ist also Teil des Stabilitätsgebietes. Befinden sich in der linken Halbebene keine Pole von R, dann ist R holomorph auf $\mathbb{C}^-$. Nach dem Maximumprinzip (Satz 2.6, S. 10) kann R in $\mathbb{C}^-$ kein

lokales Maximum haben. Stattdessen muss die Funktion ihren höchsten Betrag in der linken Halbebene auf dem Rand von $\mathbb{C}^-$ haben. Da die imaginäre Achse der Rand von $\mathbb{C}^-$ ist, kann R auf der linken Halbebene an keinem Punkt den Wert 1 überschreiten. Damit folgt A-Stabilität.
Für die andere Beweisrichtung gilt: Ist $R(z)$ A-stabil, so ist $|R(z)| \leq 1$ auf $\mathbb{C}^-$. Damit können hier keine Polstellen existieren und die imaginäre Achse kann nicht vom Ordnungsstern geschnitten werden. □

Als Beispiel für einen Ordnungsstern eines A-stabilen Verfahrens kann die Padé(3,5)-Approximation auf der vorhergehenden Seite betrachtet werden. Es ist zu sehen, dass sich alle fünf Pole auf der rechten Halbebene befinden und der Ordnungsstern die imaginäre Achse nicht schneidet.

Satz 4.4 *Wenn $R(z)$ eine Approximation der Ordnung p an e^z ist, d.h. wenn Gleichung (3.12) gilt, dann verhält sich A um den Ursprung herum wie ein „Stern“ mit $p+1$ Sektoren mit gleichem Mittelpunktswinkel $\frac{\pi}{p+1}$, getrennt durch $p+1$ Sektoren der komplementären Menge, die ebenfalls die Mittelpunktswinkel $\frac{\pi}{p+1}$ haben. Die positive reelle Achse ist um den Ursprung herum genau dann in einem Sektor von A, wenn $C < 0$ und genau dann in einem Sektor der Komplementärmenge, wenn $C > 0$.*

Beweis: Wird Gleichung (3.12) mit $e^{-z} = 1 - z + \frac{z^2}{2} - \frac{z^3}{3} + ...$ multipliziert, so ergibt sich:

$$\frac{R(z)}{e^z} = 1 - Cz^{p+1} + \mathcal{O}(z^{p+2}), \quad z \to 0 \tag{4.4}$$

Für $z \in A$ muss gelten

$$\left|\frac{R(z)}{e^z}\right| > 1. \tag{4.5}$$

Mit der Schreibweise $z = re^{i\theta}$ wird dies nun für $z \to 0$ betrachtet, d.h. für $r \to 0$. Durch Einsetzen von (4.4) in (4.5) (unter Vernachlässigung von $\mathcal{O}(z^{p+2})$) ergibt sich:

$$\begin{aligned}
& \left|1 - C(re^{i\theta})^{p+1}\right| > 1 \\
\Leftrightarrow \quad & \left|1 - Cr^{p+1}e^{i\theta(p+1)}\right| > 1 \\
\Leftrightarrow \quad & \sqrt{(1 - Cr^{p+1}\cos^2(\theta(p+1))) + (Cr^{p+1}\sin^2(\theta(p+1)))} > 1 \\
\Leftrightarrow \quad & \sqrt{1 + C^2r^{2(p+1)} - 2Cr^{p+1}\cos(\theta(p+1))} > 1 \\
\Leftrightarrow \quad & C^2r^{2(p+1)} - 2Cr^{p+1}\cos(\theta(p+1)) > 0 \\
\Leftrightarrow \quad & C\cos(\theta(p+1)) < \frac{1}{2}C^2r^{p+1} \\
\Rightarrow \quad & C\cos(\theta(p+1)) < 0
\end{aligned} \tag{4.6}$$

Die letzte Folgerung gilt für $r \to 0$, da dann der Term auf der rechten Seite verschwindet. Die Ungleichung gilt dann genau in Intervallen der Breite $\frac{\pi}{p+1}$. Die Rechnung kann analog für $\left|\frac{R(z)}{e^z}\right| \leq 1$ mit dann umgekehrten Ungleichungszeichen geführt werden. Es

gibt bei einem Umlauf von 0 bis 2π genau $2 \cdot (p+1)$ Intervalle der Breite $\frac{\pi}{p+1}$, für die abwechselnd $\left|\frac{R(z)}{e^z}\right| > 1$ und $\left|\frac{R(z)}{e^z}\right| \leq 1$.
Welche Winkelintervalle zum Ordnungsstern gehören bzw. welche Winkelintervalle zum dualen Stern gehören, kann durch Betrachten von C bestimmt werden. Für $\theta \in (-\frac{\pi}{2(p+1)}, \frac{\pi}{2(p+1)})$ ist $\cos(\theta) > 0$. Also muss für $C < 0$ dieser Sektor im Ordnungsstern und für $C > 0$ im dualen Stern enthalten sein. Es handelt sich hierbei um den Sektor, der den Anfang der positiven reellen Achse beinhaltet. □

Der Name *Ordnungsstern* bezieht sich also darauf, dass an der Anzahl der Sektoren des Sterns die Ordnung der Stabilitätsfunktion abgelesen werden kann. Hierfür können die Bilder in Abbildung 4.2 betrachtet werden. Es sind die Padé-Approximationen der Ordnung $p = k + j = 3$ dargestellt. In jedem Bild sind $4 = p + 1$ dunkelrote Sektoren um den Ursprung herum zu sehen, die alle gegen 0 hin den gleichen Mittelpunktswinkel haben (ebenso wie die grauroten Sektoren).

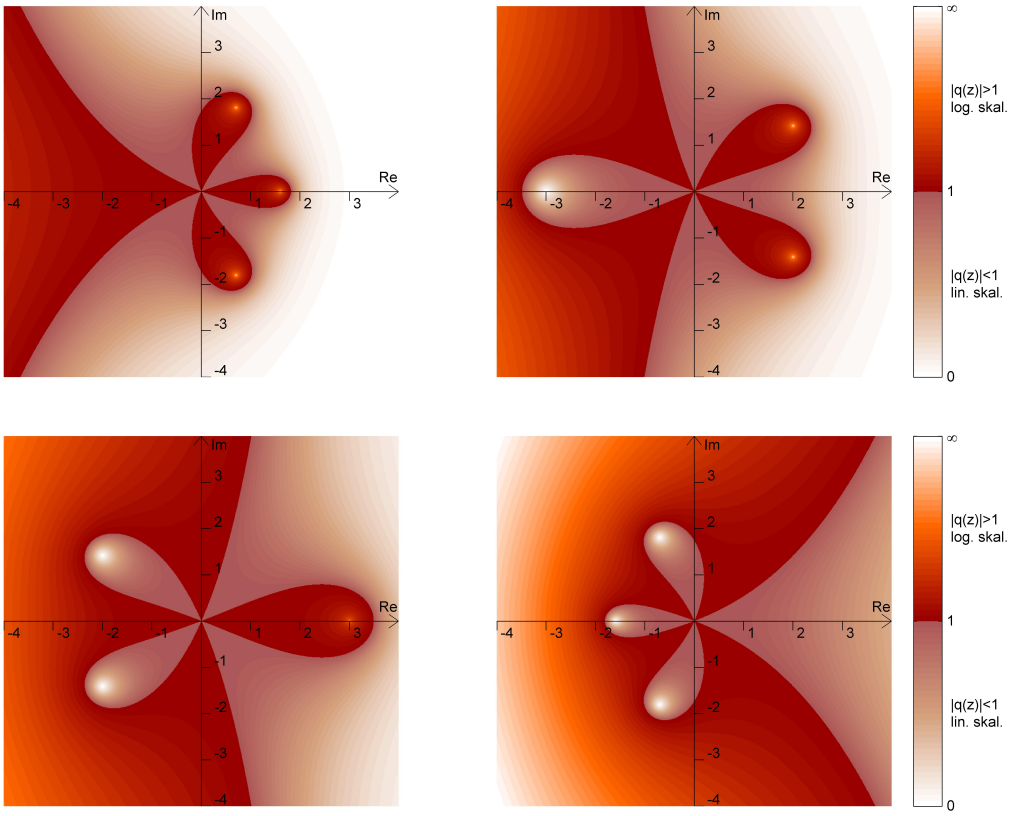

Abbildung 4.2: Ordnungssterne der Padé-Approximationen der Ordnung 3.
Oben: Padé(0,3), Padé(1,2); unten: Padé(2,1), Padé(3,0)

Es fällt auf, dass die Padé-Approximationen mit vertauschten Indices (j, k) eine Art gespiegeltes „Negativ“ darstellen.
Außerdem ist hier zu erkennen, dass es links außerhalb der beschränkten dualen Finger immer einen unbeschränkten Bereich gibt, der zum Ordnungsstern gehört, und rechts

außerhalb der Finger einen Bereich, der zum Komplement $\mathbb{C} \setminus A$ gehört. Dies wird deutlicher im folgenden Satz.

Satz 4.5 *Es gibt genau einen unbeschränkten Finger und genau einen unbeschränkten dualen Finger. Außerdem gibt es nur zwei Randlinien von ∂A die ins Unendliche gehen. Weiterhin gibt es für jedes $\varepsilon > 0$ ein $r_\varepsilon > 0$ so dass für alle $r > r_\varepsilon$ Folgendes gilt:*

$$\left\{ z = re^{i\theta} : -\frac{\pi}{2} + \varepsilon \leq \theta \leq \frac{\pi}{2} - \varepsilon \right\} \not\subseteq A \tag{4.7}$$

$$\left\{ z = re^{i\theta} : \frac{\pi}{2} + \varepsilon \leq \theta \leq \frac{3\pi}{2} - \varepsilon \right\} \subseteq A \tag{4.8}$$

Beweis: Es wird zunächst gezeigt, dass es genau zwei unbeschränkte Bereiche gibt, jeweils einen für den Ordnungsstern und für den dualen Stern. Außerdem wird der Winkelbereich dieser Finger hergeleitet. Im zweiten Teil des Beweises wird dann gezeigt, dass es nur zwei ins Unendliche führende Randlinien gibt, die diese Bereiche voneinander abgrenzen.

Teil 1 Es gilt:

$$e^{re^{i\theta}} = e^{r\cdot(\cos(\theta)+i\,\sin(\theta))} = e^{r\,\cos(\theta)} \cdot (\cos(r\,\sin(\theta)) + i\,\sin(r\,\sin(\theta))) \tag{4.9}$$

Damit folgt für den Betrag:

$$|e^{re^{i\theta}}| = e^{r\cos(\theta)} \cdot \sqrt{\cos^2(r\,\sin(\theta)) + \sin^2(r\,\sin(\theta))} = e^{r\cos(\theta)} \tag{4.10}$$

Weiterhin gilt

$$\cos(\theta) > 0,\ \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \qquad \cos(\theta) < 0,\ \theta \in \left(\frac{\pi}{2}, \frac{3\pi}{2}\right). \tag{4.11}$$

Für $\theta = \frac{\pi}{2}$ oder $\theta = \frac{3\pi}{2}$ ist dagegen $\cos(\theta) = 0$.
Sei $\varepsilon > 0$. Dann wird r_ε so groß gewählt, dass der Wert $r_\varepsilon \cos(\theta)$ für $-\frac{\pi}{2}+\varepsilon \leq \theta \leq \frac{\pi}{2}-\varepsilon$ groß genug und für $\frac{\pi}{2} + \varepsilon \leq \theta \leq \frac{3\pi}{2} - \varepsilon$ klein genug ist, damit für alle $r > r_\varepsilon$ die Exponentialfunktion bereits dominiert, d.h. dass ihr Wert bei positivem Exponenten deutlich größer und bei negativem Exponenten deutlich kleiner ist als jede rationale Funktion. Für genügend große $r > r_\varepsilon$ gilt dann

$$\theta \in \left(-\frac{\pi}{2} + \varepsilon, \frac{\pi}{2} - \varepsilon\right) \qquad \Rightarrow \frac{|R(re^{i\theta})|}{e^{r\cos(\theta)}} = |R(re^{i\theta})|e^{-r\cos(\theta)} < 1 \tag{4.12}$$

$$\theta \in \left(\frac{\pi}{2} + \varepsilon, \frac{3\pi}{2} - \varepsilon\right) \qquad \Rightarrow |R(re^{i\theta})|e^{-r\cos(\theta)} > 1 \tag{4.13}$$

Teil 2 Aus dem ersten Teil des Beweises folgt, dass es mindestens zwei unendliche Randlinien geben muss. Es bleibt also noch zu beweisen, dass der Rand von A nur genau zwei solche Linien hat. Sei wieder r_ε so groß, dass sich $z = re^{i\theta}$ für $r > r_\varepsilon$ und alle Werte von θ außerhalb von allen Null- und Polstellen von $R(z)$ befindet und die

Exponentialfunktion dominiert. Sei $r > r_\varepsilon$ fest. Es wird gezeigt, dass auf dem Kreis $z = re^{i\theta}$ für nur zwei Werte von θ gilt, dass $|R\left(re^{i\theta}\right)| = |e^{re^{i\theta}}|$. Der Einfachheit halber werden die Quadrate verglichen. Es werden also die Nullstellen der folgenden Funktion gesucht:

$$\Psi(\theta) := |e^{re^{i\theta}}|^2 - |R\left(re^{i\theta}\right)|^2 \tag{4.14}$$

Mit Gleichung (4.10) und da $|z|^2 = z \cdot \overline{z}$ ist dies äquivalent zu:

$$e^{2r\cos(\theta)} - R\left(re^{i\theta}\right) \cdot R\left(re^{-i\theta}\right) = 0 \tag{4.15}$$

Anstatt die Nullstellen hiervon explizit zu finden, wird der Mittelwertsatz angewendet. Wird $\theta = 0$ bzw. $\theta = 2\pi$ gesetzt, so ergibt sich, da r genügend groß gewählt wurde:

$$\Psi(0) = \Psi(2\pi) = e^{2r} - (R(r))^2 > 0 \tag{4.16}$$

Ebenso ergibt sich mit $\theta = \pi$:

$$\Psi(\pi) = e^{-2r} - (R(-r))^2 < 0 \tag{4.17}$$

Es wird nun die Ableitung betrachtet, um zu zeigen, dass Ψ jeweils für $\theta \in (0, \pi)$ und für $\theta \in (\pi, 2\pi)$ streng monoton fallend bzw. wachsend ist, so dass es in der oberen und unteren Halbebene jeweils genau einen Wert geben muss, für den $\Psi = 0$.
Der erste Term ergibt abgeleitet:

$$\frac{\partial}{\partial\theta}\left(e^{2r\cos(\theta)}\right) = -2r\sin(\theta)e^{2r\cos(\theta)} \tag{4.18}$$

Für den zweiten Term wird mit $R = R(re^{i\theta})$ und $\bar{R} = R(re^{-i\theta})$ abgekürzt. Es ergibt sich mit $z + \overline{z} = 2\,\mathrm{Re}(z)$:

$$\begin{aligned}
\frac{\partial}{\partial\theta}\left(R\bar{R}\right) &= R' \cdot rie^{i\theta} \cdot \bar{R} + R \cdot \bar{R}' \cdot r(-i)e^{-i\theta} \\
&= R' \cdot rie^{i\theta} \cdot \bar{R} \cdot \frac{R}{R} + R \cdot \bar{R}' \cdot r(-i)e^{-i\theta} \cdot \frac{\bar{R}}{\bar{R}} \\
&= rR\bar{R}\left(\frac{R'}{R} \cdot ie^{i\theta} + \frac{\bar{R}'}{\bar{R}} \cdot (-i)e^{-i\theta}\right) \\
&= r|R|^2 2\,\mathrm{Re}\left(ie^{i\theta}\frac{R'}{R}\right) \\
&= 2r|R\left(re^{i\theta}\right)|^2\,\mathrm{Re}\left(ie^{i\theta}\frac{R'\left(re^{i\theta}\right)}{R\left(re^{i\theta}\right)}\right)
\end{aligned} \tag{4.19}$$

Wird $R(z) = \frac{P(z)}{Q(z)}$ gesetzt, wobei P und Q Polynome sind, so ergibt sich

$$\lim_{z\to\infty}\frac{R'(z)}{R(z)} = \lim_{z\to\infty}\left(\frac{P'(z)}{P(z)} - \frac{Q'(z)}{Q(z)}\right) = 0, \tag{4.20}$$

da die Ableitung eines Polynoms geringeren Grad hat als das Polynom selbst. Also verschwindet der Term $\frac{R'}{R}$, da r genügend groß gewählt wurde, so dass als Ableitung von Ψ nur die Ableitung des ersten Terms bleibt, die in Gleichung (4.18) gegeben ist. Da $\sin(\theta)$ größer 0 ist für $\theta \in (0, \pi)$ und kleiner 0 für $\theta \in (\pi, 2\pi)$, gilt also für die Ableitung von Ψ:

$$\begin{aligned} \frac{\partial}{\partial\theta}\Psi(\theta) &< 0, \quad \theta \in (0, \pi) \\ \frac{\partial}{\partial\theta}\Psi(\theta) &> 0, \quad \theta \in (\pi, 2\pi) \end{aligned} \tag{4.21}$$

Also muss es nach dem Mittelwertsatz jeweils genau einen Schnittpunkt in der oberen Halbebene und einen in der unteren geben und somit jeweils genau eine Randlinie, die ins Unendliche geht. □

Satz 4.6 *Falls für* $R(z)$

$$R(z) = Kz^\ell + \mathcal{O}(z^{\ell-1}), \quad z \to \infty \tag{4.22}$$

mit einem $K \in \mathbb{R}$ *und* $\ell \in \mathbb{Z}$ *gilt, nähern sich die zwei unbegrenzten Randlinien des Ordnungssterns asymptotisch an*

$$x = \log(|K|) + \ell \log(|y|) \tag{4.23}$$

an.

Beweis: Auf dem Rand von A ist $|q(z)| = 1$, d.h. $|R(z)| = |e^z|$. Wird in Gleichung (4.10) statt der Schreibweise $z = re^{i\theta}$ die Schreibweise $z = x + iy$ verwendet, so ergibt sich $|e^z| = e^x$. Für $z \to \infty$ wird die gegebene Formel eingesetzt, wobei der Term $\mathcal{O}(z^{\ell-1})$ vernachlässigt wird. Dann gilt:

$$|K| \left(\sqrt{(x^2+y^2)}\right)^\ell = e^x \Rightarrow x = \log\left(|K|\right) + \ell \ \log\left(\sqrt{(x^2+y^2)}\right) \tag{4.24}$$

Aus dem vorhergehenden Satz folgt, dass die zwei ins Unendliche führenden Randlinien zwischen dem unendlichen Finger und dem unendlichen dualen Finger nicht in den beiden durch die Gleichungen (4.7) und (4.8) definierten Bereichen liegen können. Also gilt, dass am Rand von A für $x + iy \to \infty$ der Wert von x deutlich kleiner sein muss als der von y. Damit gilt $\frac{x}{y} \to 0$ und dadurch folgt:

$$x^2 + y^2 = y^2 \left(\left(\frac{x}{y}\right)^2 + 1\right) \to y^2 \tag{4.25}$$

Damit ergibt sich die gegebene Formel. □

Mit der Herleitung in Abschnitt 2.3 können für alle hier behandelten Funktionen K und ℓ und damit die Asymptoten bestimmt werden. Für explizite Verfahren sind K und ℓ sogar direkt ablesbar als der höchste Exponent (ℓ) und der Vorfaktor (K). Beim expliziten Euler-Verfahren gilt z.B. $K = 1$ und $\ell = 1$ und bei RK4 ist $K = \frac{1}{24}$ und $\ell = 4$.

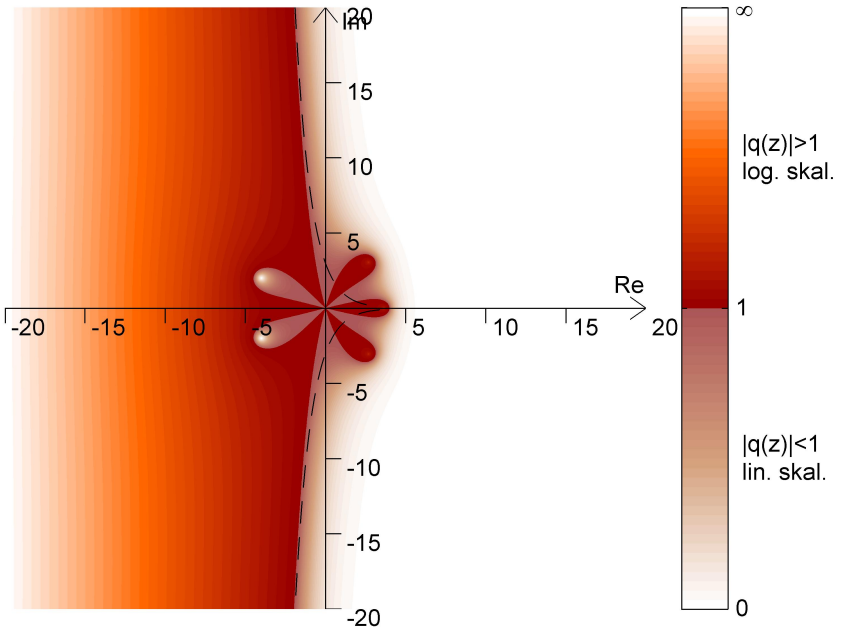

Abbildung 4.3: Ordnungsstern Padé(2,3) mit Asymptoten für ∂A

Für die Padé(2,3) Approximation ergeben sich mit Gleichung (2.33) die Werte $K = 3$ und $\ell = -1$. Sie ist in Bild 4.3 mit den Asymptoten dargestellt.
Der folgende Satz zeigt, dass sich in den Fingern die Pole von $R(z)$ befinden und in den dualen Fingern die Nullstellen.

Satz 4.7 *Jeder beschränkte Finger $F \subset A$ der Vielfachheit m beinhaltet mindestens m Pole von $q(z)$ (gezählt mit ihrer Vielfachheit) und jeder beschränkte duale Finger $F \subset \mathbb{C} \setminus A$ der Vielfachheit m beinhaltet mindestens m Nullstellen von $R(z)$.*

Beweis: Der Rand ∂F eines Fingers der Vielfachheit m besteht aus m vom Ursprung ausgehenden und zum Ursprung zurückkehrenden Randlinien. Eine solche Linie sei parametrisiert durch eine positiv orientierte Kurve $c(t)$, $t_0 \leq t \leq t_1$. Dann ist der Vektor $\vec{a} := (\dot{c}_1(t), \dot{c}_2(t))$ der Tangentenvektor zu c und $\vec{n} := (\dot{c}_2(t), -\dot{c}_1(t))$ ein nach außen gerichteter Normalenvektor.
Es wird nun analog zu der Schreibweise $z = re^{i\theta}$ die komplexwertige Funktion $q(z)$ (siehe Gleichung (4.2)) mit $z = x + iy$ aufgeteilt:

$$q(z) = r(x,y) \cdot e^{i\varphi(x,y)}, \quad \begin{aligned} r(x,y) &= |q(z)| \\ \varphi(x,y) &= \arg(q(z)) \end{aligned} \tag{4.26}$$

Die Funktion $r(x,y)$ ist also der Betrag von q und $\varphi(x,y)$ ist der Winkel. Nun wird der Logarithmus dieser Funktion betrachtet:

$$\log(q(z)) = \log(r(x,y)) + i\varphi(x,y) \tag{4.27}$$

Es gelten die Cauchy-Riemannschen Differentialgleichungen aus (2.15). Sie ergeben hier

$$\frac{\partial \log(r)}{\partial x} = \frac{\partial \varphi}{\partial y}; \quad \frac{\partial \log(r)}{\partial y} = -\frac{\partial \varphi}{\partial x}. \tag{4.28}$$

Für den Betrag von $q(z)$ gilt nach der Definition der Ordnungssterne, Def. 4.1, dass $|q(z)| > 1$ innerhalb, $|q(z)| = 1$ auf dem Rand und $|q(z)| \leq 1$ außerhalb des Ordnungssterns ist. Die Funktion $r(x, y)$ ist also konstant auf dem Verlauf von c, so dass

$$\frac{\partial \log(r)}{\partial \vec{a}} = 0 \tag{4.29}$$

gilt. Außerdem nimmt $r(x, y)$ und damit auch $\log(r)$ nach außen hin ab. Damit ergibt sich:

$$\begin{aligned} 0 \geq \frac{\partial \log(r)}{\partial \vec{n}} &= \frac{\partial \log(r)}{\partial x} \cdot n_1 + \frac{\partial \log(r)}{\partial y} \cdot n_2 \\ &= \frac{\partial \log(r)}{\partial x} \cdot \dot{c}_2(t) - \frac{\partial \log(r)}{\partial y} \dot{c}_1(t) \\ &= \frac{\partial \varphi}{\partial y} \dot{c}_2(t) + \frac{\partial \varphi}{\partial x} \dot{c}_1(t) \\ &= \frac{\partial \varphi}{\partial y} a_2 + \frac{\partial \varphi}{\partial x} a_1 = \frac{\partial \varphi}{\partial \vec{a}} \end{aligned} \tag{4.30}$$

Also nimmt der Winkel beim Durchlaufen von c ab (vergleiche die in Bild 4.4 am roten Sektoren zu sehende negative Drehung des Winkels, wenn die äußere Randlinie im positiven Sinne durchlaufen wird).

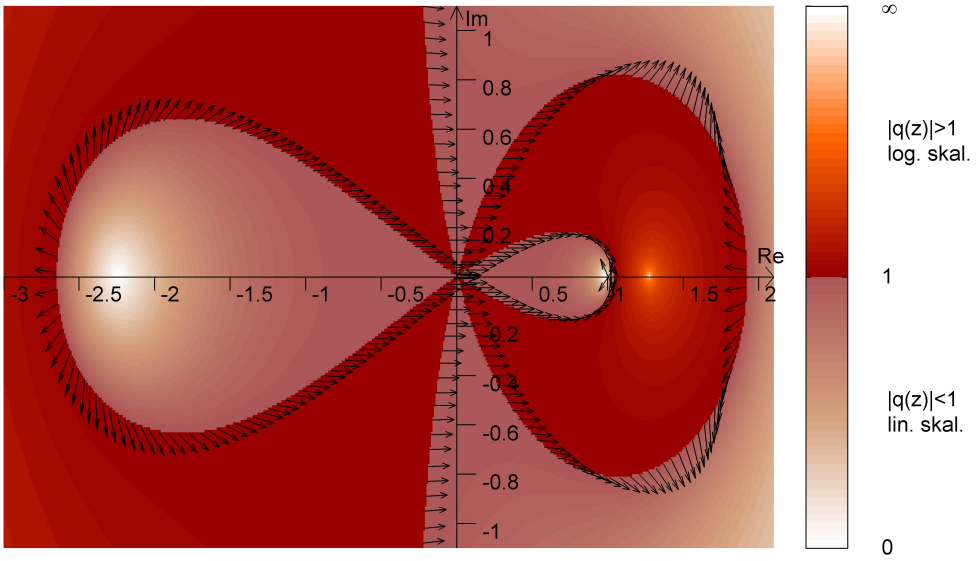

Abbildung 4.4: Ordnungsstern SDIRK3 mit φ auf ∂A eingezeichnet ($\gamma = \frac{3+\sqrt{3}}{6}$)

Um zu zeigen, dass der Winkel bis auf endlich viele Punkte sogar streng monoton fallend ist, wird die Ableitung nach t von $q(c(t))$ betrachtet. Dabei wird in der folgenden Rechnung $r = r(c_1(t), c_2(t))$ abgekürzt, analog φ und q.

$$\begin{aligned}\frac{d}{dt}q(c(t)) &= \left(\frac{\partial r}{\partial x}\,\dot{c}_1 + \frac{\partial r}{\partial y}\,\dot{c}_2\right) e^{i\varphi} + i\left(\frac{\partial \varphi}{\partial x}\,\dot{c}_1 + \frac{\partial \varphi}{\partial y}\,\dot{c}_2\right) r e^{i\varphi} \\ &= \frac{\partial r}{\partial \vec{a}}\, e^{i\varphi} + i\, q\, \frac{\partial \varphi}{\partial \vec{a}} \\ &= i\, q\, \frac{\partial \varphi}{\partial \vec{a}}\end{aligned} \tag{4.31}$$

Hierfür wurde Gleichung (4.29) benutzt.
Da die Ableitung von $q(z)$ nur eine endliche Anzahl Nullstellen hat, muss dies mit der eben berechneten Gleichung auch für $\frac{\partial \varphi}{\partial \vec{a}}$ gelten. Bis auf endlich viele Punkte ist der Winkel also streng monoton fallend bei Durchlaufen von $c(t)$.
Bei $z = 0$ ist der Winkel immer ein Vielfaches von 2π, denn $q(0) = \frac{R(0)}{e^0} \in \mathbb{R}$ und daraus folgt mit Gleichung (4.26), dass $\varphi(0,0) = 2\pi n, n \in \mathbb{Z}$. Der Winkel macht also entlang der Kurve ganze Drehungen. Da φ in Richtung der Kurve sinkt, sind diese Drehungen mit dem Uhrzeigersinn orientiert (vergleiche Bild 4.5: Auf dem Rand des beschränkten Fingers findet eine doppelte Drehung statt).

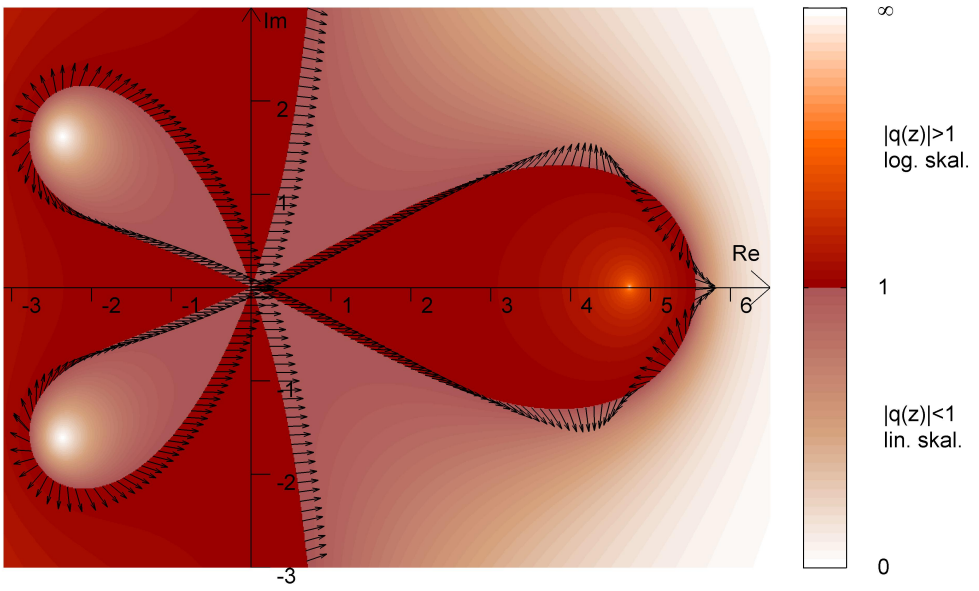

Abbildung 4.5: Ordnungsstern SDIRK3 mit φ auf ∂A eingezeichnet ($\gamma = \frac{3-\sqrt{3}}{6}$)

Für die berechnete Ableitung kann auch die Kettenregel aus Gleichung (2.3) eingesetzt werden:

$$\frac{d}{dt}q(c(t)) = q'(c(t)) \cdot \dot{c}(t) \tag{4.32}$$

Durch Umstellen von Gleichung (4.31) folgt dann:

$$\begin{aligned} & \frac{q'(c(t))}{q(c(t))} \cdot \dot{c}(t) = i \frac{\partial \varphi}{\partial \vec{a}} \\ \Rightarrow \quad & \int_{t_0}^{t_1} \frac{q'(c(t))}{q(c(t))} \cdot \dot{c}(t) dt = i \int_{t_0}^{t_1} \frac{\partial \varphi}{\partial \vec{a}} dt \\ \Leftrightarrow \quad & \oint_c \frac{q'(z)}{q(z)} dz = i \left(\varphi(c(t_1)) - \varphi(c(t_0))\right) \end{aligned} \tag{4.33}$$

Da m solche Kurven den Rand von F beschließen und φ entlang der Kurven sinkt und in jeder Kurve mindestens eine negative Volldrehung vollführt, ergibt sich also für die Differenz auf der rechten Seite:

$$\varphi(c(t_1)) - \varphi(c(t_0)) \leq -2\pi m \tag{4.34}$$

Andererseits ist $q(z)$ eine holomorphe Funktion bis auf endlich viele Punkte, die im Innern von F liegen, und es kann das Argumentenprinzip aus Satz 2.5 angewandt werden. Da keine Nullstellen innerhalb des Ordnungssterns liegen können, fällt das N im Satz weg und es gilt:

$$\oint_c \frac{q'(z)}{q(z)} dz = -2\pi i P \tag{4.35}$$

Dabei ist P die Zahl der Pole von $R(z)$ bzw. $q(z)$, die sich innerhalb dieses Fingers befinden. Somit gilt:

$$-P = \frac{1}{2\pi i} \oint_c \frac{q'(z)}{q(z)} dz \leq -m \tag{4.36}$$

Also ist $m \leq P$ und es gibt mindestens m Pole in einem Finger der Vielfachheit m.

Für die dualen Finger verläuft der Beweis analog, mit dem Unterschied, dass

$$\frac{\partial \log(r)}{\partial \vec{n}} \geq 0 \tag{4.37}$$

und somit der Winkel andersherum rotiert und sich im Inneren keine Pole sondern Nullstellen befinden. □

Als Beispiele für die Ergebnisse dieses Satzes dienen die Bilder 4.4 und 4.5 auf den vorhergehenden Seiten. Die Stabilitätsfunktion des SDIRK3-Verfahrens ist in Gleichung (3.11) auf Seite 17 gegeben.
Sie hat zwei Nullstellen und einen doppelten Pol. Bei positivem Vorzeichen von der Wurzel in γ werden zwei Sektoren um den Ursprung herum zu einem Finger, der eben diesen zweifachen Pol enthält, dies ist ein Beispiel für $m = P$. Bei negativem Vorzeichen in der Wurzel ist dagegen für das SDIRK3-Verfahren $m < P$. Der Finger, der nur die Vielfachheit $m = 1$ hat, enthält den doppelten Pol.
Allerdings ist in Abbildung 4.5 zu sehen, dass der durch Pfeile in Richtung von $q(z)$ eingezeichnete Winkel zwei ganze Drehungen vollführt und beim Schnittpunkt mit der

x-Achse wieder ein Vielfaches von 2π ist. Auf dem Rand von A ist $|q(z)| = 1$ und für $\varphi(x, y) = 0$ oder Vielfache von 2π gilt dann an diesem Punkt, dass

$$\frac{R(z)}{e^z} = r(x, y) = |q(z)| = 1 \tag{4.38}$$

und damit $R(z) = e^z$.
Eine alternative Schreibweise des Satzes ist demnach:

Satz 4.8 *Jeder beschränkte Finger $F \subset A$ mit $\partial F \subset \partial A$ hat in seinem Inneren genau so viele Pole, wie er Punkte auf dem Rand hat, für die $R(z) = e^z$ gilt. Ebenso hat jeder beschränkte duale Finger in seinem Inneren genau so viele Nullstellen, wie er Punkte auf dem Rand hat, für die die Stabilitätsfunktion mit der Exponentialfunktion übereinstimmt.*

4.2 Spezielle Ordnungssterne

Bei den meisten Verfahren und Stabilitätsfunktionen, die in dieser Arbeit betrachtet werden, gibt es nur die aus den Sektoren hervorgehenden Finger. Es gibt aber auch Sonderfälle der Approximationen an die Exponentialfunktion, die z.B. Finger haben, die nicht mit dem Ursprung verbunden sind.
Wenn diesen Fingern analog zur Definition 4.2 die Vielfachheit 0 zugeordnet wird, so ergeben sich keine Widersprüche zu allen bislang genannten Ergebnissen.

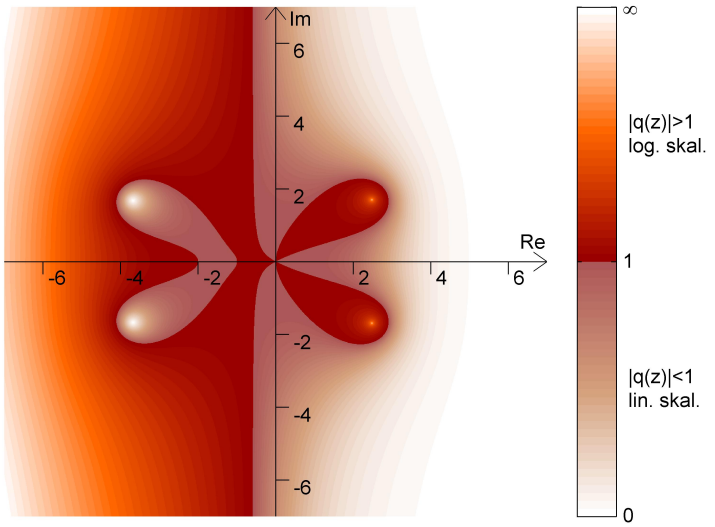

Abbildung 4.6: Ordnungsstern einer 2/2-Approximation der Ordnung 2

Als Beispiel zeigt Abbildung 4.6 den Ordnungsstern einer von S.P. Nørsett in [13] eingeführten Funktion, die einen dualen Finger besitzt, der nicht mit dem Ursprung

verbunden ist. Es handelt sich um eine beschränkten Padé-Approximation mit

$$R(z) = \frac{1 + \frac{2-\alpha}{4}z + \frac{1-\alpha+\beta}{8}z^2}{1 - \frac{2+\alpha}{4}z + \frac{1+\alpha+\beta}{8}z^2}, \tag{4.39}$$

wobei α und β so gewählt wurden, dass $R(-1) = e^{-1}$ und $R(-2) = e^{-2}$. Die Funktion wird von Nørsett die 2/2-Approximation der Ordnung 2 genannt. Die Ordnung kann an den drei roten Sektoren um den Ursprung herum wie üblich abgelesen werden.
Die in diesem Kapitel dargestellten Sätze gelten genauso auch für diese Art Funktionen. Das Verhalten in den Fingern der Vielfachheit 0 ändert also nichts an den in dieser Arbeit vorgestellten Ergebnissen. Die zweite Version von Satz 4.7, Satz 4.8, kann sogar sehr gut auf den nicht mit dem Ursprung verbundenen dualen Finger in Nørsetts Approximation angewendet werden. Die Stabilitätsfunktion ist so gewählt, dass sie an diesem dualen Finger an zwei Stellen mit der Exponentialfunktion übereinstimmt. Daher müssen in diesem Bereich zwei Nullstellen liegen.

4.3 Beweis von Ehles Vermutung mit Hilfe von Ordnungssternen

Wird noch einmal Satz 4.4 für Ordnungssterne zu rationalen Stabilitätsfunktionen betrachtet, so können auch durch Betrachten der Ordnungssternbilder leicht einige Vermutungen zur A-Stabilität von Padé-Approximationen aufgestellt werden.

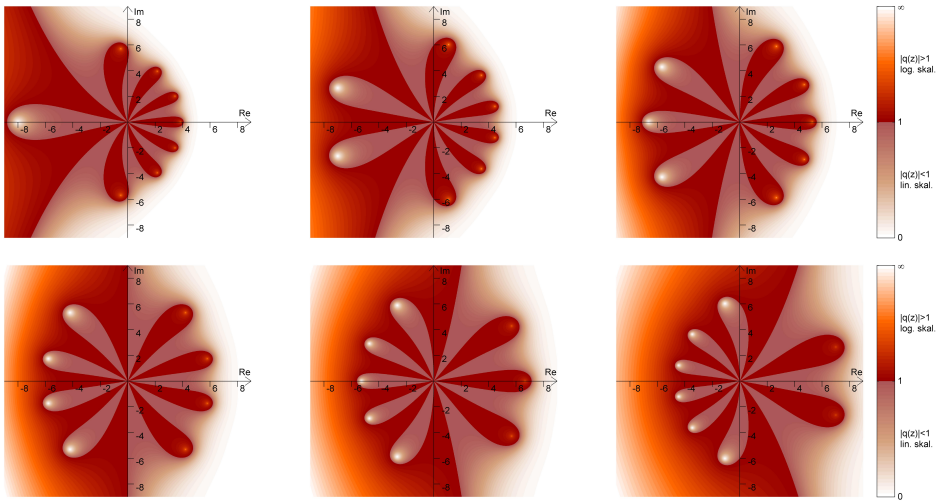

Abbildung 4.7: Ordnungssterne. oben: Padé(1,7), Padé(2,6) Padé(3,5);
unten: Padé(4,4), Padé(5,3), Padé(6,2)

Dies wird unterstrichen durch die Bilderreihe einiger Padé-Approximationen der Ordnung 8 in Abbildung 4.7. Aus den Bildern ist klar zu erkennen, dass für $k > j$ die Funktion nicht A-stabil ist. Insgesamt ist zu sehen, dass nur die Padé(3,5) und Padé(4,4) Approximationen A-stabil sind.

Es lässt sich vermuten, dass bei Padé-Approximationen zusätzlich zu $k \leq j$ zum Erreichen von A-Stabilität $j \leq k+2$ gelten muss. Sonst reichen entweder Pole in die linke Halbebene hinüber oder die Finger, in denen diese Pole liegen, beginnen links der imaginären Achse und kreuzen die imaginäre Achse.
Ehle hat diese Vermutung bereits 1969 geäußert, in [9] gelang es 1978 Hairer, Wanner und Nørsett erstmals, sie zu beweisen.
In diesem Abschnitt wird über ein paar Sätze auf diesen Beweis hingearbeitet. Dabei ist $R(z)$ zunächst eine beliebige rationale Approximation der Ordnung p mit k Nullstellen und j Polstellen.

Satz 4.9 *Wenn $R(z)$ A-Stabil ist, dann gilt $p \leq 2k_1 + 2$, dabei ist k_1 die Anzahl der unterschiedlichen Nullstellen in $\mathbb{C}^-$.*

Beweis: Es gibt mindestens $\lfloor \frac{p+1}{2} \rfloor$ Sektoren von A, die in der linken Halbebene beginnen[3]. Da sich keine Pole in der linken Halbebene befinden, sind alle diese Sektoren zusammen nach Satz 4.3 unbeschränkt und schließen also mindestens $\lfloor \frac{p+1}{2} \rfloor - 1$ duale Finger ein, die dann nach Satz 4.7 mindestens genausoviele Nullstellen beinhalten. Also gilt $\lfloor \frac{p+1}{2} \rfloor - 1 \leq k_1$. Durch Multiplizieren mit 2 ergibt sich die Behauptung. (Ist p gerade, so ergibt $2\lfloor \frac{p+1}{2} \rfloor = p$, ist p ungerade so gilt $2\lfloor \frac{p+1}{2} \rfloor = p+1$. Da $p < p+1$ gilt die Ungleichung immer.) □

Die Voraussetzung $|R(iy)| \leq 1,\ y \in \mathbb{R}$ im folgenden Satz wird auch als I-Stabilität bezeichnet. Der Satz wird aber später unter der Voraussetzung von A-Stabilität benutzt. I-Stabilität ist ein Teil von A-Stabilität. Jedes A-stabile Verfahren ist auch I-stabil. Umgekehrt muss aber zusätzlich noch die Funktion R analytisch in $\mathbb{C}^-$ sein, es dürfen sich also keine Pole in der linken Halbebene befinden.

Satz 4.10 *Wenn $|R(iy)| \leq 1,\ \forall y \in \mathbb{R}$, dann gilt $p \leq 2j_1$, wobei j_1 die Anzahl der Pole von $R(z)$ in $\mathbb{C}^+$ sei.*

Beweis: Mindestens $\lfloor \frac{p+1}{2} \rfloor$ Sektoren von A beginnen in der rechten Halbebene. Sie können nicht in die linke Halbebene hinüber reichen und müssen daher nach Satz 4.5 beschränkt sein. Also müssen sie wiederum nach Satz 4.7 mindestens genauso viele Pole in ihrem Innern haben, so dass $\lfloor \frac{p+1}{2} \rfloor \leq j_1$ woraus die Behauptung wie oben folgt. □

Satz 4.11 *$R(z) = \frac{P(z)}{Q(z)}$ ist A-stabil, wenn folgendes gilt:*

(i) $p \geq 2j - 2$,

(ii) $\lim_{z \to \infty} |R(z)| \leq 1$ *und*

(iii) die Koeffizienten von Q besitzen alternierende Vorzeichen.

[3]Bei der Notation $\lfloor a \rfloor$ handelt es sich um die sogenannte Gauß-Klammer, die definiert ist als die größte Zahl $r \in \mathbb{Z}$, für die $r \leq a$ gilt.

Beweis: Für den Beweis wird das E-Polynom aus Definition 3.5 verwendet. Aus (ii) folgt $\deg P = k \leq j = \deg Q$ und damit mit Satz 3.6 $\deg(E(y)) \leq 2j$ mit $y \in \mathbb{R}$. Außerdem gilt mit (i)

$$p \geq 2j - 2 \Rightarrow p + 1 \geq 2j - 1. \tag{4.40}$$

Nach Satz 3.6 gilt auch noch

$$E(y) = \mathcal{O}(y^{p+1}), \quad y \to 0 \tag{4.41}$$

und dass das E-Polynom eine gerade Funktion ist. Somit ergibt sich für den Grad $\deg(E) = 2j$, also $E(y) = Ky^{2j}$ für eine Konstante K. Wird zusätzlich noch

$$(ii) \quad \Rightarrow \quad |R(z)|^2 \leq 1, \ z \to \infty \tag{4.42}$$

hinzugezogen, so ergibt sich

$$\begin{aligned} Ky^{2j} &= |Q(iy)|^2 - |P(iy)|^2 \\ \Rightarrow \qquad \lim_{y\to\infty} \frac{Ky^{2j}}{|Q(iy)|^2} &= \lim_{y\to\infty} \left(1 - |R(iy)|^2\right) \geq 0 \end{aligned} \tag{4.43}$$

und somit $K \geq 0$. Daraus folgt $E(y) \geq 0$ und damit $|R(iy)| \leq 1 \ \forall y$, d.h. $R(z)$ ist I-stabil.

Es muss also noch gezeigt werden, dass sich keine Pole in $\mathbb{C}^-$ befinden.
Wird (i) in Satz 4.10 eingesetzt, so ergibt sich $2j-2 \leq p \leq 2j_1$ also kann sich maximal ein Pol in der linken Halbebene $\mathbb{C}^-$ befinden. Dieser muss also reell sein, da komplexe Pole wie komplexe Nullstellen immer paarweise konjugiert auftreten müssen. Befindet sich ein reeller Pol von $R(z)$ in $\mathbb{C}^-$ so gibt es eine reelle negative Nullstelle von Q. Sei

$$Q(z) = q_0 + q_1 z^1 + q_2 z^2 + ... + q_j z^j. \tag{4.44}$$

Nach (iii) haben die Koeffizienten $q_0, q_1, ...q_j$ alternierende Vorzeichen. Dann gilt für $Q(-z)$, dass alle Vorzeichen der Koeffizienten gleich sind. Nach Descartes Regel über Vorzeichen[12] ist die Anzahl der negativen reellen Nullstellen höchstens so groß, wie die Anzahl der Vorzeichenwechsel von $Q(-z)$, also kann kein Pol in $\mathbb{C}^-$ existieren. □

Satz 4.12 *Wenn $R(z)$ k_0 unterschiedliche Nullstellen und j_0 unterschiedliche Pole besitzt, dann gilt $p \leq k_0 + j_0$.*

Der Beweis für diesen Satz wird mit der Graphentheorie geführt und benutzt das Eulersche Theorem, nach dem für einen zusammenhängenden planaren Graphen mit v Knoten, q Kanten und r Gebieten $v + r = q + 2$ gilt.[2] Für einen nicht zusammenhängenden Graphen, der aus g Komponenten besteht, gilt $v + r - q = 1 + g$.
Im Beweis wird auf die möglichen anderen Bereiche der Ordnungssterne (vergleiche Abschnitt 4.2) mit eingegangen. Er ist doppelt ausgeführt, zunächst nur für die „regulären“ Ordnungssterne und dann allgemeiner.

Beweis: Angenommen der Ordnungsstern besteht nur aus Zusammenhangskomponenten, die wie in den bisherigen Darstellungen mit dem Ursprung verbunden sind: Der Ordnungssterns wird als planarer Graph auf $\mathbb{C} \cup \infty$ dargestellt, so dass alle Randlinien Kanten sind. Die beiden ins Unendliche gehenden Randlinien werden zu einer Kante vom Ursprung zum Ursprung zurück vereinigt. Siehe dazu Bild 4.8 unten.
Die Knoten werden definiert als diejenigen Punkte z auf ∂A, bei denen sich zwei Randlinien schneiden. Der Ursprung ist in diesem Fall Schnittpunkt aller Kanten und daher der einzige Knoten. Die Anzahl der Gebiete sei r, die Zahl der Kanten sei q und die Zahl der Knoten ist $v = 1$. Es gilt also $r + 1 = q + 2$ (i).
Es gibt zwei „unendliche“ Gebiete (die in dieser Darstellung durch die durch Unendlich gehende Kante begrenzt sind), alle anderen Gebiete enthalten jeweils mindestens einen Pol oder eine Nullstelle. Somit ist $r \leq k_0 + j_0 + 2$ (ii).
Der Grad eines Knotens ist definiert als die Zahl der Kanten, die dort beginnen oder enden. Werden die Gradzahlen aller Knoten summiert, so wird jede Kante doppelt gezählt. Nach Satz 4.4 haben $2(p+1)$ Kanten den Ursprung als einen Endpunkt. Also muss $2p + 2 = 2q$ (iii) gelten.
Es ergibt sich also:

$$p \overset{(iii)}{=} q - 1 \overset{(i)}{=} r - 2 \overset{(ii)}{\leq} k_0 + j_0 \tag{4.45}$$

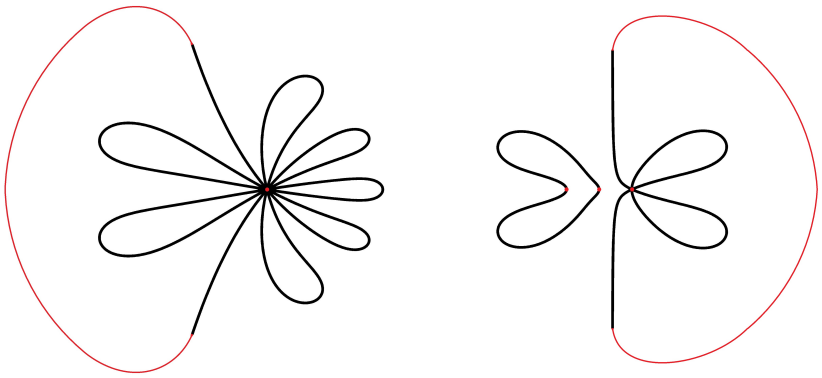

Abbildung 4.8: Ordnungsstern Graphentheorie (links: Padé(2,5); rechts: Nørsetts 2/2-Approximation der Ordnung 2)

Der Beweis wird nun verallgemeinert. Der Graph des Ordnungssterns hat q Kanten, r Gebiete und v Knoten. Es wird angenommen, dass er aus mehreren Komponenten besteht, mindestens jedoch aus einer, somit gilt $v + r - q \geq 2 \Rightarrow q - v \leq r - 2$ (i).
Die Ungleichung $r \leq k_0 + j_0 + 2$ (ii) gilt immer noch, aber die dritte Formel muss angepasst werden. Die Knoten sind wie oben definiert. An jedem Knoten enden also mindestens zwei Kanten. Für den Knoten im Ursprung gilt weiterhin, dass $2(p + 1)$ Kanten an diesem Knoten enden, für die $v-1$ anderen Knoten gilt nach der Definition, dass mindestens zwei Kanten in ihnen enden. Insgesamt ergibt die Summe über die

Gradzahlen aller Knoten $2q$ wie oben. Somit ergibt sich hier eine dritte Ungleichung $2q \geq 2(p+1) + 2(v-1) \Rightarrow p \leq q - v$ (iii).
Mit diesen drei Ungleichungen folgt:

$$p \overset{(iii)}{\leq} q - v \overset{(i)}{\leq} r - 2 \overset{(ii)}{\leq} k_0 + j_0 \tag{4.46}$$

□

Satz 4.13 (Ehles Vermutung) *Eine Padé-Approximation $R(z)$ ist A-stabil genau dann, wenn*

$$j - 2 \leq k \leq j. \tag{4.47}$$

Alle Nullstellen und Pole sind einfach.

Beweis: Bei Padé-Approximationen gilt $p = k+j$. Mit der Gleichung aus Satz 4.12 und da $k_0 \leq k$ und $j_0 \leq j$ folgt $p \leq k_0 + j_0 \leq k + j = p$, also muss überall Gleichheit herrschen und $(k - k_0) + (j - j_0) = 0$. Da beide eingeklammerten Werte ≥ 0 sind, müssen beide Terme jeweils 0 ergeben. Also sind alle Nullstellen und Pole einfach.
Für die „genau dann, wenn"-Beziehung ($\Leftrightarrow$) gilt für die Hinrichtung ($\Rightarrow$): Aus Satz 4.9 folgt für eine A-stabile Funktion $p \leq 2k_1 + 2 \leq 2k + 2$ also $k + j \leq 2k + 2$ also $j - 2 \leq k$. Aus Satz 4.10 wiederum folgt $p \leq 2j_1 \leq 2j$ woraus $k + j \leq 2j \Leftrightarrow k \leq j$ folgt.
Die Rückrichtung folgt aus Satz 4.11, denn mit $p = j + k$ folgt aus $j - 2 \leq k$ die Ungleichung $2j - 2 \leq j + k = p$, mit $k \leq j$ folgt $\lim_{z\to\infty} |R(z)| \leq 1$, da der Nenner größeren Grad hat als der Zähler und nach der Definition von Padé-Approximationen sind die Koeffizienten von Q alternierend. □

4.4 Relative Ordnungssterne

Stehen mehrere Lösungsverfahren zur Auswahl, kann es sinnvoll sein, nicht jede einzeln mit der Exponentialfunktion zu vergleichen. Zum Vergleich untereinander wurde die Definition des relativen Ordnungssterns eingeführt.

Definition 4.14 *Seien $R_1(z)$ und $R_2(z)$ zwei rationale Approximationen an e^z. Dann ist ihr relativer Ordnungsstern definiert als das Gebiet*

$$B = \left\{ z \in \mathbb{C} : \left| \frac{R_1(z)}{R_2(z)} \right| > 1 \right\}. \tag{4.48}$$

Die für die Ordnungssterne und für Stabilitätsfunktionen im Allgemeinen eingeführten Begriffe lassen sich auch auf die relativen Ordnungssterne übertragen.
So ist die Ordnung der Approximation gegeben durch $p = \min(p_1, p_2)$, wobei p_1 die Ordnung von $R_1(z)$ ist und p_2 die Ordnung von $R_2(z)$ und es gilt:

$$\frac{R_1(z)}{R_2(z)} = 1 - Cz^{p+1} + \mathcal{O}(z^{p+2}), \quad z \to 0 \tag{4.49}$$

Das kann hergeleitet werden, indem Gleichung (3.12) für beide Funktionen angewendet wird und dann voneinander abgezogen wird:

$$\begin{aligned} e^z - R_1(z) &= C_1 z^{p_1+1} + \mathcal{O}(z^{p_1+2}), \quad z \to 0 \\ e^z - R_2(z) &= C_2 z^{p_2+1} + \mathcal{O}(z^{p_2+2}), \quad z \to 0 \\ \Rightarrow R_1(z) - R_2(z) &= C_2 z^{p_2+1} - C_1 z^{p_1+1} + \mathcal{O}(z^{\min(p_1,p_2)+2}), \quad z \to 0 \end{aligned} \tag{4.50}$$

Für den Term auf der rechten Seite gilt:

$$C_2 z^{p_2+1} - C_1 z^{p_1+1} = \begin{cases} C_2 z^{p_2+1}\left(1 - \frac{C_1}{C_2} z^{p_1-p_2}\right), & p_1 > p_2 \\ -C_1 z^{p_1+1}\left(-\frac{C_2}{C_1} z^{p_2-p_1} + 1\right), & p_2 > p_1 \\ (C_2 - C_1) z^{p_1+1}, & p_1 = p_2 \end{cases} \tag{4.51}$$

Der Koeffizient C ergibt sich also als $C = -C_2$ für $p_1 > p_2$ bzw. als $C = C_1$ für $p_2 > p_1$ oder als $C = C_1 - C_2$ für den Fall $p_1 = p_2$. Es wird davon ausgegangen, dass $C \neq 0$, da sonst identische Verfahren verglichen werden. Dann folgt mit $p = \min(p_1, p_2)$, dass

$$R_1(z) - R_2(z) = -C z^{p+1} + \mathcal{O}(z^{p+2}), \quad z \to 0. \tag{4.52}$$

Es wird nun durch $R_2(z)$ geteilt bzw. mit $(R_2(z))^{-1} = \frac{Q_2(z)}{P_2(z)}$ multipliziert. Dabei handelt sich wieder um eine rationale Funktion, die sich in eine Potenzreihe umwandeln lässt. Auf der rechten Seite ergibt sich dann

$$\begin{aligned} C z^{p+1} \cdot \frac{Q_2(z)}{P_2(z)} + \mathcal{O}(z^{p+2}) &= C z^{p+1} \cdot (1 + a_1 z + a_2 z^2 + ...) + \mathcal{O}(z^{p+2}),\ z \to 0 \\ &= C z^{p+1} + \mathcal{O}(z^{p+2}),\ z \to 0, \end{aligned} \tag{4.53}$$

wobei die a_i die Koeffizienten der Potenzreihe von $(R_2(z))^{-1}$ seien. Der erste Term der Potenzreihe ist die 1, da $R_2(z)$ eine Approximation an die Exponentialfunktion der Ordnung $p_2 \geq 0$ ist und somit der erste Term der Potenzreihe von R_2 nach Gleichung (3.12) mit dem ersten Term der Potenzreihe der Exponentialfunktion übereinstimmen muss. Dann ist auch der erste Term der Potenzreihe von R_2^{-1} gleich 1. Durch Addieren von 1 auf beiden Seiten ergibt sich dann Gleichung (4.49).
In Bild 4.9 werden Beispiele relativer Ordnungssterne gezeigt. Es werden das explizite und implizite Euler-Verfahren verglichen, wobei hier zu sehen ist, dass die Abbildung für B mit $|\frac{R_1(z)}{R_2(z)}|$ von den Wertebereichen genau umgekehrt ist zum Vergleich für $|\frac{R_2(z)}{R_1(z)}|$. Außerdem wird noch Runge-Kutta-4 mit Runge-Kutta-3 und SDIRK3 mit $\gamma = \frac{3-\sqrt{3}}{6}$ zu SDIRK3 mit $\gamma = \frac{3+\sqrt{3}}{6}$ verglichen.
Die Eigenschaften von Ordnungssternen lassen sich größtenteils auch auf relative Ordnungssterne übertragen. So bleibt die Aussage von Satz 4.4 erhalten, was sich leicht zeigen lässt, indem statt Gleichung (4.4) die hier verwendete Gleichung (4.49) benutzt wird.
Da für den Vergleich einer rationalen Funktion mit einer anderen das Verhalten für $z \to \infty$ nicht so eindeutig vergleichbar ist wie mit der Exponentialfunktion, da beide

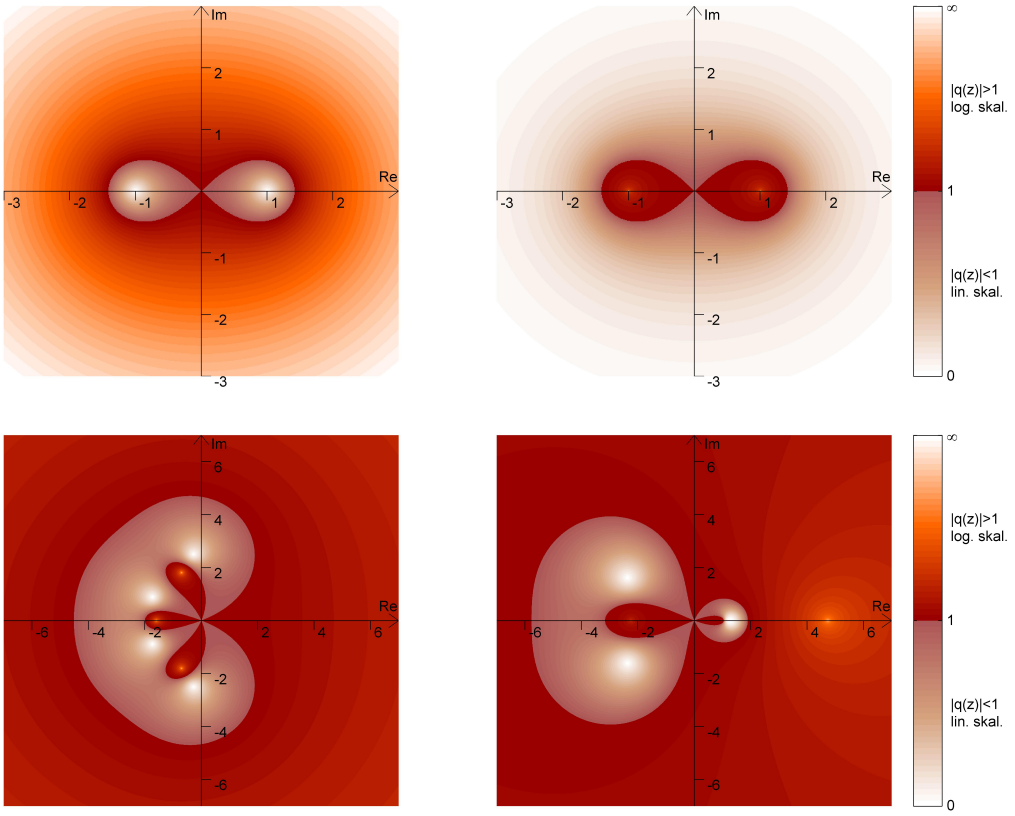

Abbildung 4.9: Relative Ordnungssterne. Oben links: Explizites zu implizitem Euler-Verfahren; rechts: imp. zu exp. E.-V.; unten links: RK4 zu RK3, rechts: SDIRK3($\gamma = \frac{3-\sqrt{3}}{6}$) zu SDIRK3($\gamma = \frac{3+\sqrt{3}}{6}$)

Funktionen Pole und Nullstellen haben, kann Satz 4.5 nicht für relative Ordnungssterne angewandt werden.
Dagegen gilt Satz 4.7 auch für die unbeschränkten Finger von B, denn während für A immer ∞ als Pol nur durch $e^{-z} \to \infty$ erreicht werden kann, muss ∞ als Pol des Bruchs aus rationalen Funktionen auch beachtet werden. Auch hier wurde im Beweis die Exponentialfunktion nicht verwendet und kann somit ersetzt werden.
Wäre die Exponentialfunktion die Stabilitätsfunktion eines Verfahrens, so wäre ihr Stabilitätsgebiet, das heißt das Gebiet, auf dem ihr Betrag ≤ 1 ist, die linke Halbebene $\mathbb{C}^-$ und der Rand ihres Stabilitätsgebietes die imaginäre Achse, da dort gilt $|e^{iy}| = 1$. Wenn also eine Methode mit Stabilitätsfunktion $R(z)$ A-stabil ist, so heißt das im Vergleich mit der Exponentialfunktion als zweiter Stabilitätsfunktion, dass ihr Stabilitätsgebiet das von e^z beinhaltet.
Auf diese Art ergeben sich die folgenden Analogien zwischen einem Ordnungsstern A und einem relativen Ordnungsstern B, dabei bezeichnet S_1 das Stabilitätsgebiet von

$R_1(z)$ und S_2 das von $R_2(z)$.

$$\begin{aligned} \text{Ordnungsstern } A &\leftrightarrow \text{relativer Ordnungsstern } B \\ \text{imaginäre Achse} &\leftrightarrow \partial\, S_2 \\ \mathbb{C}^- &\leftrightarrow \text{Innengebiet von } S_2 \\ \mathbb{C}^+ &\leftrightarrow \text{Außengebiet von } S_2 = \mathbb{C} \setminus S_2 \\ \text{Verfahren ist A-stabil} &\leftrightarrow S_1 \supset S_2 \end{aligned} \tag{4.54}$$

Werden mit diesen Analogien in den Sätzen 4.9 und 4.10 die Begriffe für relative Ordnungssterne eingesetzt, so ergeben sich zwei weitere Sätze, die sich auch auf relative Ordnungssterne übertragen lassen. Letzterer wird im Beweis des folgenden Satzes gebraucht, daher wird er an dieser Stelle für relative Ordnungssterne umformuliert:

Satz 4.15 *Gilt $|R_1(z)| \leq 1$ für alle $z \in \partial S_2$, dann ist $p \leq 2j_1$, wobei j_1 die Anzahl der Pole von $\frac{R_1(z)}{R_2(z)}$ im Bereich $\mathbb{C} \setminus S_2$ sei.*

Der Beweis ist analog zum Beweis im Ordnungssternfall, nur dass für Satz 4.7 nicht die Beschränktheit der Finger gelten muss und somit nicht auf Satz 4.5 eingegangen werden muss.
Es wird jetzt ein kleines Theorem zum Vergleich von zwei Stabilitätsfunktionen vom Grad s (also zum Beispiel zwei expliziten Runge-Kutta-Verfahren mit der gleichen Anzahl Stufen) angegeben.

Satz 4.16 *Seien die Stabilitätsfunktionen $R_1(z)$ und $R_2(z)$ Polynome vom Grad s und Ordnungen ≥ 1, dann gilt für die zugehörigen Stabilitätsbereiche:*

$$S_1 \not\supset S_2 \qquad \textit{und} \qquad S_1 \not\subset S_2 \tag{4.55}$$

Beweis: Wäre $S_1 \supset S_2$, dann wäre $|R_1(z)| \leq 1$ für alle $z \in \partial\, S_2$. Also muss mit dem vorhergehenden Satz und da $p \geq 1$ mindestens ein Pol von $\frac{R_1(z)}{R_2(z)}$ außerhalb von S_2 liegen. Da $deg(R_1) = deg(R_2)$ kann es keinen Pol bei $z = \infty$ geben. Also wären die einzigen möglichen Pole die Nullstellen von S_2, diese liegen aber in S_2 und damit ergibt sich ein Widerspruch. Es muss also gelten $S_1 \not\supset S_2$. Durch Austauschen von R_1 und R_2 ergibt sich analog $S_1 \not\subset S_2$. □

Dieses Ergebnis zeigt, dass keine von beiden Methoden immer besser sein kann, es gibt also immer eine Differentialgleichung, die nur von der ersten und immer eine Differentialgleichung, die nur von der zweiten Methode ausreichend gut gelöst werden kann. Der nächste Satz (Jeltsch und Nevanlinna, 1981, vgl. [11]) verallgemeinert diese Aussage für Funktionen von unterschiedlichem Grad.
Der Satz vergleicht Verfahren mit verschiedenen Anzahlen an Stufen, also Verfahren, die nicht nur möglicherweise unterschiedliche Ordnung haben, sondern die auch unterschiedlich viel Rechenaufwand betreiben. Um solche Verfahren vergleichbar zu machen, wird der Begriff der skalierten Stabilitätsgebiete eingeführt:

Definition 4.17 *Sei $R(z)$ die Stabilitätsfunktion vom Grad s eines expliziten Runge-Kutta-Verfahrens mit s Stufen, dann heißt*

$$S^{scal} = \{z : |R(sz)| \leq 1\} = \{z : sz \in S\} = \frac{1}{s}S \tag{4.56}$$

skalierter Stabilitätsbereich *der Methode.*

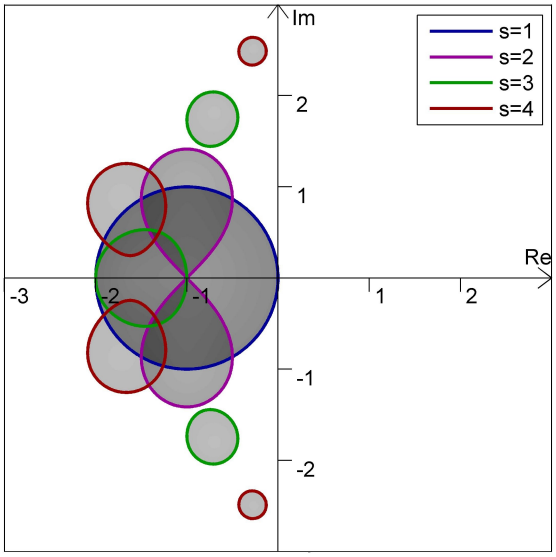

Abbildung 4.10: Skalierte Stabilitätsbereiche expliziter Runge-Kutta-Verfahren

Im Bild 4.10 sind die skalierten Stabilitätsbereiche der Runge-Kutta-Verfahren 1-4 gezeigt. Zusätzlich veranschaulicht diese Abbildung das Ergebnis des nun folgenden Satzes, der mit Hilfe von Ordnungssternen 1981 erstmals bewiesen werden konnte.

Satz 4.18 *Seien $R_1(z)$ und $R_2(z)$ die Stabilitätsfunktionen zweier expliziter Runge-Kutta-Verfahren mit $deg(R_1) = s_1$ und $deg(R_2) = s_2$. Wenn beide Verfahren eine Ordnung ≥ 1 haben, dann kann keiner ihrer skalierten Stabilitätsbereiche den anderen komplett in seinem Inneren haben, d.h. es gilt:*

$$S_1^{scal} \not\supset S_2^{scal} \quad und \quad S_2^{scal} \not\supset S_1^{scal} \tag{4.57}$$

Beweis: Um die numerische Arbeit beider Methoden vergleichen zu können, werden vom ersten Verfahren s_2 Schritte mit Schrittweite $\frac{h}{s_2}$ ausgeführt und vom zweiten Verfahren s_1 Schritte mit Schrittweite $\frac{h}{s_1}$. Wird noch einmal die Definition der Stabilitätsfunktion in Definition 3.1 und die Herleitung davon betrachtet, dann ist erkennbar, dass unter Anwendung der Dahlquist'schen Testgleichung nach der besagten Zahl Schritte hier die Bereiche

$$\left\{z \in \mathbb{C} : \left|\left(R_1\left(\frac{z}{s_2}\right)\right)^{s_2}\right| \leq 1\right\} = \left\{z \in \mathbb{C} : \left|\left(R_1\left(\frac{z}{s_2}\right)\right)\right| \leq 1\right\} = \left\{z \in \mathbb{C} : \frac{1}{s_2}z \in S_1\right\} = s_2 \cdot S_1 \tag{4.58}$$

und analog

$$\left\{ z \in \mathbb{C} : \left| \left(R_2 \left(\frac{z}{s_1} \right) \right)^{s_1} \right| \leq 1 \right\} = s_1 \cdot S_2 \tag{4.59}$$

verglichen werden. Die Funktionen

$$\tilde{R}_1 = \left(R_1 \left(\frac{z}{s_2} \right) \right)^{s_2} \quad \text{und} \quad \tilde{R}_2 = \left(R_2 \left(\frac{z}{s_1} \right) \right)^{s_1} \tag{4.60}$$

sind vom selben Grad $s_1 s_2$. Mit dem Satz 4.16 gilt dann

$$s_2 \cdot S_1 \not\supset s_1 \cdot S_2 \Rightarrow S_1^{scal} \not\supset S_2^{scal} \tag{4.61}$$

und analog $S_1^{scal} \not\subset S_2^{scal}$. □

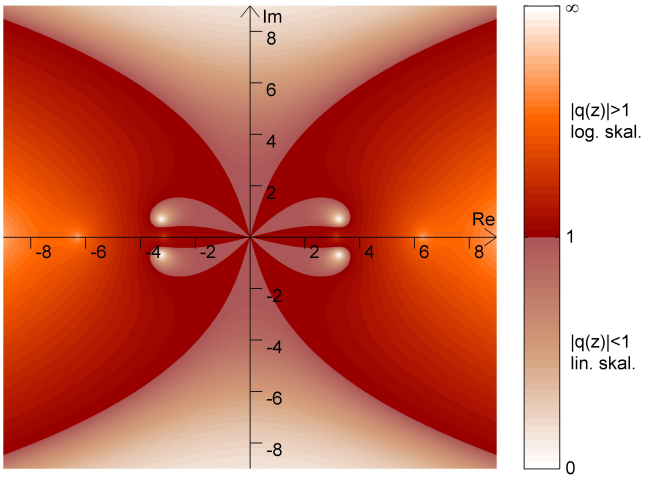

Abbildung 4.11: Ordnungsstern einer Approximation an $\sin(z)$

In dieser Arbeit werden Ordnungssterne nur im Rahmen der Analyse von Stabilität von Verfahren zum Lösen von Differentialgleichungen dargestellt. In diesen Abschnitt wird deutlich, dass mit einem Ordnungsstern nicht immer eine Approximation der Exponentialfunktion betrachtet werden muss. Es können, völlig losgelöst vom Thema der numerischen Lösung von Differentialgleichung, Ordnungssterne betrachtet werden, um Approximationen an beliebige Funktionen zu untersuchen. Nørsett und Iserles gehen in [13] verstärkt auf diese Art Ordnungssterne ein.

Als Beispiel wird für die Sinusfunktion eine Approximation gewählt, indem die Potenzreihe des Sinus nach den ersten drei Termen abgebrochen wird:

$$R(z) = z + \frac{1}{3!} z^3 - \frac{1}{5!} z^5 \tag{4.62}$$

Wird nun $\tilde{q}(z) = \frac{R(z)}{\sin(z)}$ gesetzt, so ergibt sich der in Bild 4.11 dargestellten Ordnungsstern. Dadurch, dass die Sinusfunktion periodisch Nullstellen hat, sind auf der x-Achse periodisch Pole zu sehen.

4.5 Reelle Pole

Viele bekannte und wichtige Verfahren zum Lösen von (steifen) Differentialgleichungen haben reelle Pole. Dazu gehören zum Beispiel die DIRK-Methoden und damit auch die SDIRK-Verfahren.
Für die diagonal impliziten Runge-Kutta-Verfahren lässt sich z.B. mit Gleichung (3.10) durch Einsetzen der Koeffizienten aus (2.8) mit $a_{i,i} = \gamma_i$ und $a_{i,j} = 0$ für $i < j$ mit der folgenden Gleichung für die Stabilitätsfunktion zeigen, dass die Pole alle reell sind (siehe auch Gleichung (3.11)):

$$R(z) = \frac{P(z)}{(1-\gamma_1 z)(1-\gamma_2 z)...(1-\gamma_s z)} \tag{4.63}$$

Reelle Pole schränken die Ordnung eines Verfahrens ein. Sind reelle Pole vorhanden, so wird zu sehen sein, dass $p-1$ nur noch durch die Zahl der Nullstellen k und die komplexen Pole beschränkt ist. Ist gar kein komplexer Pol vorhanden, wird die Ordnung umso mehr beschränkt.
Der folgende Satz wurde in [9] so zum ersten Mal bewiesen.

Satz 4.19 *Sei $R(z) = \frac{P(z)}{Q(z)}$ gegeben mit $deg(P) = k$, $deg(Q) = j$ und von den j Nullstellen seien nur m unterschiedliche Nullstellen komplexwertig. Besitzt $Q(z)$ zusätzlich reelle Nullstellen, dann gilt für die Ordnung:*

$$p \leq k + m + 1 \tag{4.64}$$

Hat $Q(z)$ keine reellen Nullstellen, dann gilt sogar:

$$p \leq k + m \tag{4.65}$$

Beweis: Wenn es m komplexe unterschiedliche Pole gibt, so gibt es maximal m Finger, die diese enthalten. Gibt es auch noch reelle Pole, so geht höchstens ein Finger vom Ursprung aus, innerhalb dessen äußerem Rand sich mehrere Pole (und Nullstellen) befinden können. Das liegt an der Symmetrie bezüglich der x-Achse. Es gibt dann also höchstens $m+1$ „äußere“ Finger, um die herum sich nur der unendliche Bereich von $\mathbb{C} \setminus A$ befindet. Wird auch noch der unendliche Sektor von A dazu gezählt, so gibt es also höchstens $m+2$ Gebiete, die abgewechselt werden müssen durch $m+2$ ins Unendliche gehende Sektoren des dualen Sterns.
Da es $p+1$ duale Sektoren gibt, gilt also, dass mindestens $p+1-(m+2)$ der dualen Sektoren beschränkt sind. Also gilt nach Satz 4.7 für die Nullstellen $k \geq p+m-1$ und daraus folgt die Behauptung. Gibt es keine reellen Pole, so gibt es nur noch maximal m Finger, um die herum sich nur noch der unendliche Bereich von $\mathbb{C} \setminus A$ befindet. Somit gibt es nur maximal $m+1$ ins Unendliche gehende Sektoren vom dualen Stern. □

Abbildung 4.12 zeigt den Ordnungsstern einer Stabilitätsfunktion mit einem einfachen reellen und zwei doppelten komplexen Polstellen. Die Anzahl der unterschiedlichen komplexen Pole ist $m = 2$, die Anzahl der Nullstellen ist $k = 2$ und die Ordnung ist

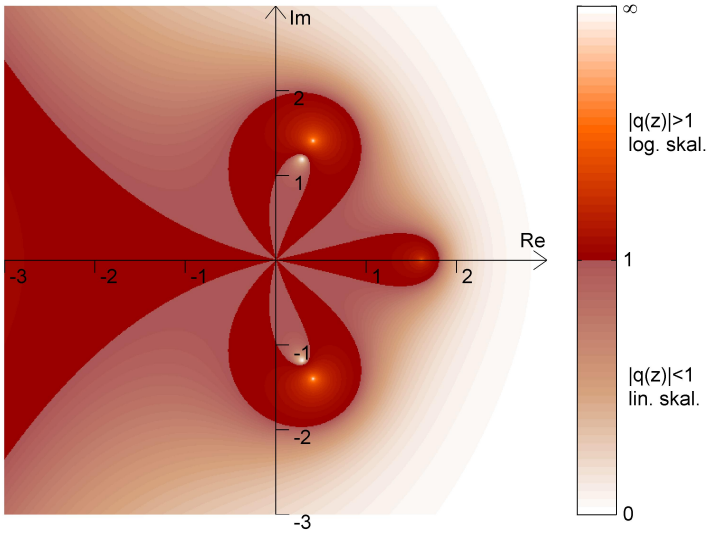

Abbildung 4.12: Ordnungsstern einer Funktion mit 2 komplexen unterschiedlichen Polstellen

maximal mit $p = 5 = k + m + 1$. Die $m + 2 = 4$ ins Unendliche gehenden Gebiete des dualen Sterns sind sehr gut zu erkennen.

Für eine Stabilitätsfunktion die gar keine komplexen Polstellen hat, wie z.B. die DIRK-Verfahren, muss als Folgerung aus Satz 4.19 gelten:

Folgerung 4.20 *Sei* $R(z) = \frac{P(z)}{Q(z)}$. $Q(z)$ *besitze nur reelle Nullstellen. Dann gilt:*

$$p \leq k + 1 \tag{4.66}$$

Wenn für eine solche Funktion die optimale Ordnung $p = k + 1$ erreicht wird, dann gibt es $p + 1 = k + 2$ Sektoren des Ordnungssterns und ebensoviele duale Sektoren. Davon gehen genau 2 ins Unendliche und somit muss dann gelten, dass die übrigen k Sektoren alle genau eine Nullstelle von R enthalten.
Hieraus kann auch noch mit Satz 4.4 gefolgert werden, dass $C > 0$, wenn es eine positive reelle Nullstelle gibt, die kleiner ist als die Polstellen, und $C < 0$ sonst.

4.6 Ordnungssterne bei Mehrschrittverfahren

Zwei weitere Vermutungen zur A-Stabilität können mit Hilfe von Ordnungssternen bewiesen werden. Bei diesen Vermutungen geht es um Mehrschrittverfahren. Die zweite Dahlquist-Schranke sagt, dass kein A-stabiles Mehrschrittverfahren eine Ordnung größer als 2 haben kann und die Daniel-Moore-Vermutung drückt als Verallgemeinerung davon aus, dass kein A-stabiles s-stufiges Runge-Kutta-Mehrschrittverfahren eine Ordnung größer als $2s$ haben kann. Diese Schranke wurde auch in [9] mit Hilfe von Ordnungssternen bewiesen.

Auch für Mehrschrittverfahren können Ordnungssterne dargestellt werden. Bei den Stabilitätsbereichen von Mehrschrittverfahren (siehe Abschnitt 3.6) wurde bereits das charakteristische Polynom erklärt. Die Stabilitätsbereiche werden in der z-Ebene angezeigt. Das charakteristische Polynom lässt sich leicht nach z umstellen. Wird nun andersherum $\zeta(z)$ ausgewertet, so gibt es typischerweise mehr als eine Wurzel $\zeta(z)$.
Zur besseren Erklärung wird wieder ein Beispiel verwendet. Das BDF-Verfahren der Ordnung 2 hat das charakteristische Polynom:

$$\left(\frac{3}{2} - z\right)\zeta^2 - 2\zeta + \frac{1}{2} = 0 \tag{4.67}$$

Die Nullstellen $\zeta(z)$ sind dann:

$$\zeta_{1,2} = \frac{2 \pm \sqrt{1+2z}}{3-2z} \tag{4.68}$$

Es ergibt sich eine Riemann'sche Fläche, die beiden Lösungsblätter ζ_1 und ζ_2 können an der reellen Achse für $x \leq -\frac{1}{2}$ verbunden werden. Sie haben alleine dort jeweils einen Diskontinuitätsschnitt.

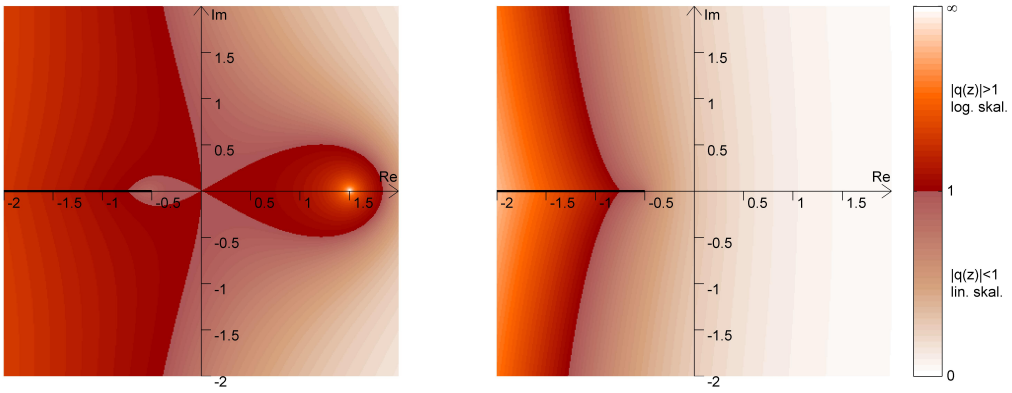

Abbildung 4.13: Ordnungssternblätter des BDF2-Verfahrens. Links ζ_1, rechts ζ_2

Ebenso ergeben sich für die Ordnungssterne bei Mehrschrittverfahren Riemann'sche Flächen. Dabei werden die einzelnen Wurzeln mit der Exponentialfunktion verglichen und die einzelnen Blätter des Ordnungssterns sind definiert als

$$A_j = \{z \in \mathbb{C} : |\zeta_j(z)| > |e^z|\}\,. \tag{4.69}$$

Die Abbildung 4.13 zeigt die beiden Blätter des BDF-Verfahrens der Ordnung 2.
Es lässt sich dann ein Hauptblatt definieren, auf dem an einem Ordnungsstern wie bei den Einschrittverfahren die Ordnung abgelesen werden kann, im Bild ist dieses links zu sehen und die Ordnung zwei an den drei roten Sektoren abzulesen.
Für weitere Details und die Beweise der angesprochenen Vermutungen sei auf die Literatur verwiesen.

5 Ordnungspfeile

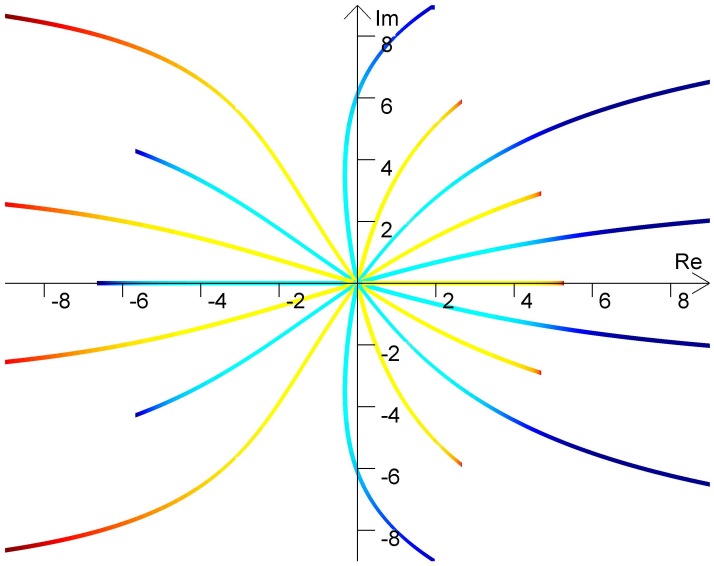

Abbildung 5.1: Ordnungspfeile Padé(3,5)

5.1 Eigenschaften von Ordnungspfeilen

Die Funktion $q(z) = \frac{R(z)}{e^z}$ bildet von $\mathbb{C}$ in $\mathbb{C}$ ab. Im Prinzip wäre hier also ein vierdimensionales Bild nötig, um alle Eigenschaften der Funktion ausreichend darzustellen. In Kapitel 4 ging es um Ordnungssterne, bei denen nur der Betrag von q betrachtet wird. Als anderen möglichen Wert, der betrachtet werden kann, gibt es noch den Winkel bzw. das Argument von q. In Abbildung 5.2 links sind die Winkel von $q(z)$ für die Padé(1,2)-Approximation dargestellt, rechts sind nur die Konturlinien einiger Winkel angezeigt.
John Butcher hat die Ordnungspfeile eingeführt als alternative Methode, die Stabilität von einem Verfahren zu analysieren. Es sollten nur die Linien betrachtet werden, die vom Ursprung aus die höchste Änderung im Betrag, positiv oder negativ, haben. Wenn $q(z)$ wieder wie in Gleichung (4.26) geschrieben wird und nach r abgeleitet wird, so ergibt sich:

$$\frac{\partial}{\partial r}\left(r(x,y)e^{i\varphi(x,y)}\right) = e^{i\varphi(x,y)} = \cos(\varphi(x,y)) + i\sin(\varphi(x,y)) \tag{5.1}$$

Diese Ableitung wird maximiert, wenn $\varphi = 0$ und dann ist $q(x+iy) = r(x,y) \geq 0$.

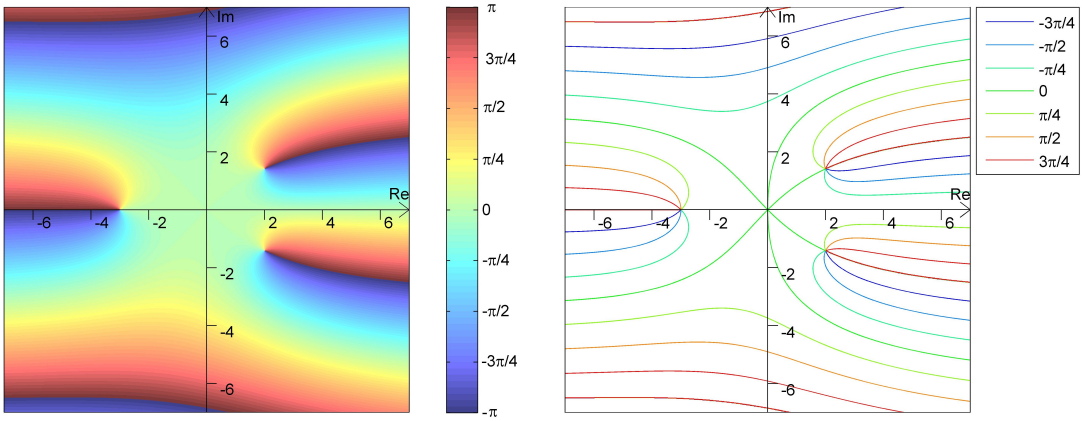

Abbildung 5.2: Winkel von $q(z)$ für die Padé(1,2)-Approximation

Bei Ordnungspfeilen werden genau die Winkelkonturlinien von $q(z)$, auf denen das Argument 0 ist und der Realteil positiv, betrachtet. Es werden also von den Konturlinien im Bild 5.2 rechts jeweils nur diejenigen verwendet, die den Winkel 0 darstellen und bei denen $q \geq 0$ ist. Auf diesen Linien kann jetzt noch der Betrag von q dargestellt werden.
Bei Butcher wird das klassisch mit Pfeilen gemacht, die in Richtung steigenden Betrags zeigen. In dieser Arbeit werden Farben verwendet, dabei zeigen Farbverläufe von gelb nach rot aufsteigenden Wert für Werte > 1 und von hellblau nach dunkelblau abfallenden Wert von 1 bis 0.

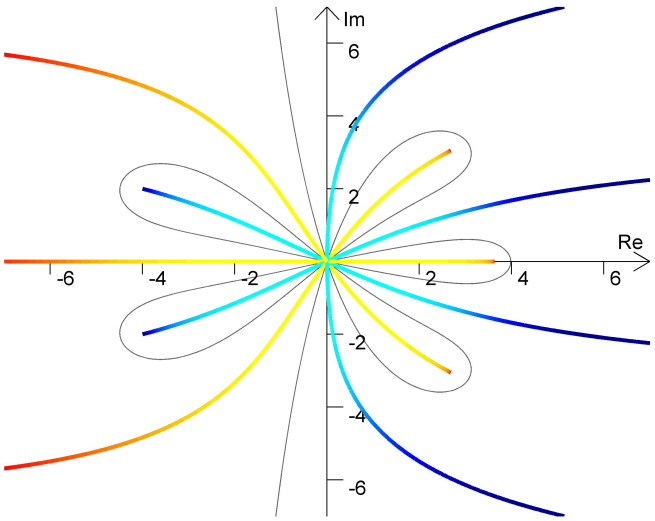

Abbildung 5.3: Ordnungspfeile mit Ordnungssternkontur (Padé(2,3)-Approximation)

Als Beispiel ist am Anfang dieses Kapitels bereits das Bild 5.1 gegeben, dass wieder die Padé(3,5)-Approximation zeigt. Bild 5.3 zeigt einmal für die Padé(2,3)-Approximation

Ordnungspfeile im Vergleich mit dem Ordnungsstern, dessen Randlinien eingezeichnet sind.

Definition 5.1 *Die Menge der Punkte in $\mathbb{C}$ für die $q(z) = \frac{R(z)}{e^z}$ positiv und reell ist, wird das* ***Ordnungsnetz*** *der rationalen Funktion $R(z)$ genannt. Die Linien, die vom Ursprung ausgehen mit wachsendem Wert von q, heißen* ***Aufwärtspfeile****, diejenigen für die q vom Ursprung aus sinkt,* ***Abwärtspfeile****. Alle Pfeile, die direkt vom Ursprung ausgehen, werden zusammengefasst als das* ***Hauptordnungsnetz****.*

Die Abtrennung des Hauptordnungsnetzes von den restlichen Ordnungspfeilen ergibt sich daraus, dass für die Betrachtungen zur Stabilität und die Analyse der Verfahren nur die Ordnungspfeile betrachtet werden, die zum Hauptordnungsnetz dazu gehören.

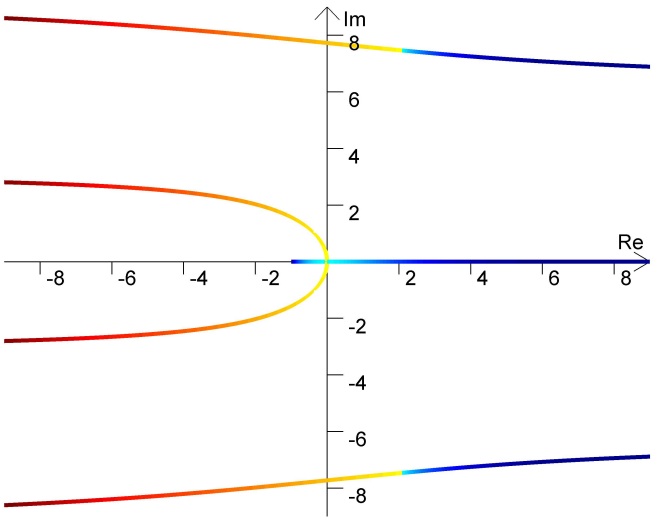

Abbildung 5.4: Ordnungspfeile explizites Euler-Verfahren

Bei den Ordnungssternen gibt es Sterne, bei denen alle Sektoren vom Ursprung ausgehen und es keine anderen Gebiete gibt. Bei den Ordnungspfeilen hat jedes Ordnungsnetz Linien, für die die Bedingungen der Ordnungspfeile zutreffen, die aber nicht vom Ursprung ausgehen (siehe als Beispiel Bild 5.4).
Das liegt an den Eigenschaften der Exponentialfunktion. Wenn z gegen ∞ geht, dann wird $q(z)$ von der Exponentialfunktion dominiert. In Richtung der Imaginärachse ist das Verhalten der Exponentialfunktion periodisch, da

$$e^{x+iy} = e^x(\cos(y) + i\sin(y)) \tag{5.2}$$

und somit für festes x das Verhalten der Sinus- bzw. Kosinusfunktion angenommen wird.
Für jede rationale Funktion $R(z)$ gibt es für $z \to \infty$ Ordnungspfeile, die näherungsweise parallel zur x-Achse von $+\infty$ nach $-\infty$ mit steigendem Betrag von $q(z)$ verlaufen. In

Abschnitt 5.2 wird beschrieben, dass bei manchen Ordnungsnetzen auch noch weitere nicht mit dem Ursprung verbundene Linien hinzukommen.
Wie schon in den Beispielbildern zu sehen ist, sind die Ordnungspfeile um 0 systematisch angeordnet. Das wird im nächsten Satz beschrieben.

Satz 5.2 *Sei $R(z)$ eine Approximation an e^z der Ordnung p, d.h. es gilt Gleichung (3.12). Für die Fehlerkonstante gelte $C \neq 0$.*
Ist $C < 0$, so gibt es Aufwärtspfeile, die vom Ursprung ausgehen und zu denen im Ursprung die Strahlen mit Winkel $\frac{2k\pi}{p+1}$ mit $k = 0, 1, ..., p$ tangential verlaufen. Außerdem existieren Abwärtspfeile die bei 0 tangential sind zu den Strahlen mit Winkel $\frac{(2k+1)\pi}{p+1}$ mit $k = 0, 1, ..., p$.
Ist $C > 0$ so sind die Winkel der Aufwärts- und Abwärtspfeile im Ursprung genau andersherum verteilt.

Beweis: Es wird wieder $z = re^{i\theta}$ gesetzt und die Gleichung für die Ordnung umgeschrieben zu (4.4). Eingesetzt ergibt sich:

$$q(re^{i\theta}) = \frac{R(re^{i\theta})}{e^{re^{i\theta}}} = 1 - Cr^{p+1}e^{i(p+1)\theta} + \mathcal{O}(r^{p+2}) \quad r \to 0 \tag{5.3}$$

Für die Winkel $\theta = \frac{2k\pi}{p+1}$ und $\theta = \frac{(2k+1)\pi}{p+1}$ ergeben sich reelle Werte. Wird nun nach r abgeleitet und der Term $\mathcal{O}(r^{p+2})$ vernachlässigt, da r sehr klein ist, so folgt:

$$\frac{\partial}{\partial r} q(re^{i\theta}) = -Cpr^p e^{i(p+1)\theta} \tag{5.4}$$

Ist $\theta_0 = \frac{2k\pi}{p+1}$ mit $k = 0, 1, ..., p$, handelt es sich also bei $(p+q)\theta_0$ um gerade Vielfache von π, dann ist $e^{i(p+1)\theta_0} = 1$, für ungerade Vielfache $\theta_1 = \frac{(2k+1)\pi}{p+1}$ mit $k = 0, 1, ..., p$ ist $e^{i(p+1)\theta_1} = -1$. Also ergibt sich für $C > 0$ bei den Winkeln mit geraden Vielfachen, dass $\frac{\partial q}{r} < 0$ und somit, dass q abfällt, und bei denen mit ungeraden Vielfachen von π, dass $\frac{\partial q}{r} > 0$, also dass q ansteigt. Für $C < 0$ ergeben sich genau umgekehrt bei den Winkeln mit geraden Vielfachen Aufwärtspfeile und bei denen mit ungeraden Vielfachen die Abwärtspfeile. □

Dieser Satz ist das Analogon zu Satz 4.5 bei den Ordnungssternen. Bei Ordnungssternen ist die Breite eines Sektors näherungsweise $\frac{\pi}{p+1}$, bei den Ordnungspfeilen ist ihr Abstand um den Ursprung herum näherungsweise durch diesen Bruch gegeben und es gibt von jeder Sorte Pfeile jeweils $p + 1$.
Bei den Ordnungssternen existierte Symmetrie zur x-Achse, da der Betrag von q betrachtet wurde. Auch die Ordnungspfeile sind zur x-Achse symmetrisch verteilt. Wenn der Wert $q(z) = \frac{R(z)}{e^z}$ reell ist, dann kann aufgrund der Regeln zum konjugierten Wert durch Einsetzen festgestellt werden, dass $e^{\overline{z}} = \overline{e^z}$ und $R(\overline{z}) = \overline{R(z)}$ und somit $q(\overline{z}) = \overline{q(z)}$ und das ist für reelle Werte das Gleiche wie $q(z)$.
Die Abwärts- und Aufwärtspfeile wechseln sich nach dem vorhergehenden Satz immer ab. Es wurde bereits beobachtet, dass einige zu den Nullstellen und Polstellen führen, das wird mit dem nächsten Satz bewiesen.

Satz 5.3 *Die Aufwärtspfeile führen entweder zu Polstellen von $R(z)$ oder zu $-\infty$. Die Abwärtspfeile führen zu den Nullstellen von $R(z)$ oder zu $+\infty$.*

Beweis: Es wird zunächst ähnlich vorgegangen wie beim Beweis von Satz 4.7. Es gilt also wieder:

$$q(z) = r(x,y) \cdot e^{i\varphi(x,y)}, \quad \begin{aligned} r(x,y) &= |q(z)| \\ \varphi(x,y) &= \arg(q(z)) \end{aligned} \tag{5.5}$$

Ein vom Ursprung ausgehender Aufwärtspfeil wird parametrisiert durch eine Kurve $c(t) = c_1(t) + ic_2(t)$ und der Tangentenvektor ist gegeben durch $\vec{a} = (\dot{c}_1(t), \dot{c}_2(t))$. Der Winkel φ ist entlang des Pfeils konstant 0, also ist auch die Ableitung $\frac{\partial \varphi}{\partial \vec{a}} = 0$. Da $\varphi = 0$, ist außerdem:

$$q(c(t)) = r(c_1(t), c_2(t)) \tag{5.6}$$

Für den Betrag $r(c_1(t), c_2(t))$ gilt entlang eines Aufwärtspfeiles, dass er wächst, also ist $\frac{\partial r}{\partial \vec{a}} \geq 0$. Diese Ungleichung ist nur in endlich vielen Fällen nicht strikt. Das lässt sich wieder durch die Ableitung von q nach t zeigen:

$$\frac{d}{dt} q(c(t)) = \frac{\partial\, r(c_1(t), c_2(t))}{\partial x}\, \dot{c}_1(t) + \frac{\partial\, r(c_1(t), c_2(t))}{\partial y}\, \dot{c}_2(t) = \frac{\partial\, r(c_1(t), c_2(t))}{\partial \vec{a}} \tag{5.7}$$

Die linke Seite kann wiederum umgeschrieben werden zu:

$$q'(c(t))\, \dot{c}(t) = \frac{\partial}{\partial \vec{a}} r(c_1(t), c_2(t)) \tag{5.8}$$

Da die Ableitung von c nicht 0 werden kann, sind also die einzigen möglichen Stellen, an denen r entlang des Pfeils stagniert, solche Stellen, für die $\frac{d}{dz} q(c(t)) = 0$. Es gibt in jedem Fall nur endlich viele solcher Stellen. Diese Stellen können nur Sattelpunkte von $q(z)$ sein, da es nach dem Maximumprinzip keine Minima und Maxima geben kann. Solche Sattelpunkte müssen immer mindestens zwei Aufwärtspfeile haben, die zu ihnen hinführen, und zwei, die von ihnen wegführen.
Es wird dann die Weiterführung der Aufwärtspfeile jeweils so definiert, dass ein entweder auf der gleichen Halbebene (obere oder untere) oder auf der x-Achse weiterführender Aufwärtspfeil den hier ankommenden zugeordnet wird. Diese Art Sattelpunkt tritt z.B. an der x-Achse auf, vergleiche Bild 5.6 links. Ein Pfeil kann mit dieser Definition die x-Achse auch nicht kreuzen, denn immer wenn sich zwei Pfeile treffen, muss es sich um einen Sattelpunkt oder eine Polstelle handeln.
Also gilt, dass der Betrag immer weiter ansteigt, ein Aufwärtspfeil kann also nur in einem Pol oder bei Unendlich enden. Es wird nun gezeigt, dass nur für $z \to -\infty$ der Wert von q gegen ∞ geht.
Alle Arten kritischer Punkte (Pole, Sattelpunkte) innerhalb von $\mathbb{C}$ hängen nur mit dem Verhalten von $R(z)$ und nicht mit der Exponentialfunktion zusammen. Geht aber nun $z \to \infty$ so gibt es wegen der Dominanz der Exponentialfunktion keine weiteren kritischen Punkte.
Es wird also für eine Stelle z auf einem Aufwärtspfeil angenommen, dass $|z|$ bereits so groß ist, dass es keinen Pol oder Sattelpunkt mehr geben kann.

Für $R(z)$ wird Gleichung (2.28) angewandt. Die Schreibweisen $z = x + iy$ und $z = re^{i\theta}$ sind äquivalent, die erste wird nun in Formel (2.28) für $R(z)$ eingesetzt und die zweite in e^z. Dabei ist r nicht mehr der Betrag von q sondern die Polarkoordinate $r = \sqrt{x^2 + y^2}$. Dann gilt:

$$\begin{aligned} q(z) = R(z)e^{-z} &= \left(Kr^\ell e^{\ell i\theta} + \mathcal{O}\left(r^{\ell-1}e^{(\ell-1)i\theta}\right)\right) \cdot e^{-x-iy}, && z \to \infty \\ &= Kr^\ell e^{-x-iy} e^{\ell i\theta} \left(1 + \mathcal{O}(r^{-1}e^{-i\theta})\right), && z \to \infty \\ &= Kr^\ell e^{-x} e^{i(\ell\theta - y)} \left(1 + \mathcal{O}(r^{-1})\right), && z \to \infty \end{aligned} \tag{5.9}$$

Der Winkel θ muss mit der oben erklärten Definition des Verhaltens an der x-Achse in dem Bereich $[0, \pi]$ oder $[\pi, 2\pi]$ bleiben. Auf einem Ordnungspfeil muss $q(z)$ reell sein. Der Wert ℓ ist in $\mathbb{Z}$. Mit diesen drei Bedingungen ergibt sich, dass y nicht gegen ∞ (oder $-\infty$) gehen kann, sondern beschränkt ist. Damit also $q(z)$ gegen ∞ geht, muss $x \to -\infty$ gelten und damit ist bewiesen, dass Aufwärtspfeile zu $-\infty$ führen. Für Abwärtspfeile wird analog vorgegangen. □

Mit der letzten Gleichung im Beweis kann berechnet werden, an welchen Stellen die näherungsweise zur x-Achse parallelen Ordnungslinien die y-Achse schneiden. Für das explizite Euler-Verfahren ist $\ell = 1$ und $K = 1$. Der Schnittpunkt mit der positiven y-Achse findet bei $\theta = \pi/2$ statt und damit folgt $n\pi = (\ell\theta - y)$ mit $n \in \mathbb{Z}$ und also $y = (\frac{1}{2} - n)\pi$. Für $n = -2$ ergibt sich die in Bild 5.4 zu sehende Linie bei $y \approx 7,85$.

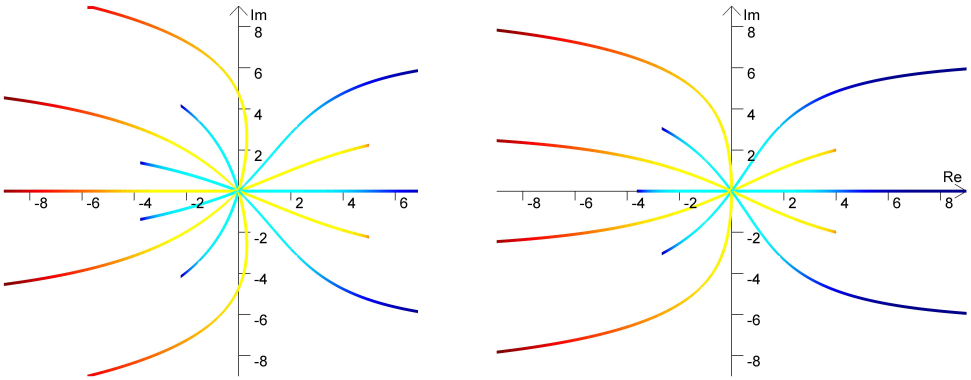

Abbildung 5.5: Ordnungspfeile. Links: Padé(4,2)-Approximation; rechts: Padé(3,2)-Approximation

Es wird jetzt das Kriterium für A-Stabilität bei Ordnungspfeilen gegeben (vergleiche A-Stabilität bei Ordnungssternen in Satz 4.3). Die beiden Bilder in Abbildung 5.5 zeigen zwei nicht A-stabile Verfahren. Im Beweis wird anhand der Ordnungspfeile erklärt, warum sie nicht A-stabil sein können.

Satz 5.4 *Ist ein Verfahren mit Stabilitätsfunktion $R(z)$ A-stabil, dann gilt Folgendes:*

(i) Alle Pole von $R(z)$ liegen in $\mathbb{C}^+$

(ii) Keiner der Aufwärtspfeile des Ordnungsnetzes schneidet die imaginäre Achse

(iii) Keiner der Aufwärtspfeile des Ordnungsnetzes ist an irgendeiner Stelle tangential zur imaginären Achse

Beweis: Es wird gezeigt, dass A-Stabilität nicht gelten kann, wenn die drei Bedingungen nicht gelten. Gäbe es Pole in der linken Halbebene, so könnte $R(z)$ nicht A-stabil sein. Wäre ein Aufwärtspfeil tangential zu der Imaginärachse oder würde er diese schneiden, so würde es ein $y \in \mathbb{R}$ geben, sodass

$$1 < \left| \frac{R(iy)}{e^{iy}} \right| = |R(iy)|, \tag{5.10}$$

da Aufwärtspfeile immer einen Wert > 1 haben. Ein Verfahren kann aber nur A-stabil sein, wenn es auch auf der Imaginärachse stabil ist und somit ist die Behauptung bewiesen. □

John Butcher stellt in [1] das Ordnungspfeilbild einer Funktion mit 6 reellen Polstellen bei $x = 4, 5, 6, 7, 8$ und 9 der Ordnung 6 vor. Die Funktionsgleichung ist dann durch (3.15) gegeben als:

$$R(z) = \frac{60480 + 264z - 5402z^2 - 719z^3 + 131z^4 + (44 + 1/6)z^5 + (3 + 7/60)z^6}{60480 - 60216z + 24574z^2 - 5265z^3 + 625z^4 - 39z^5 + z^6} \tag{5.11}$$

Es stellt sich heraus, dass diese Funktion ein sehr gutes Beispiel dafür ist, dass der Satz für A-Stabilität bei Ordnungspfeilen keine Äquivalenzbeziehung darstellt. Aus A-Stabilität folgen die drei genannten Bedingungen. Funktionen, die diese Bedingungen verletzen, sind somit nicht A-stabil. Aber Funktionen, für die alle drei Voraussetzungen gelten, müssen nicht unbedingt A-stabil sein.

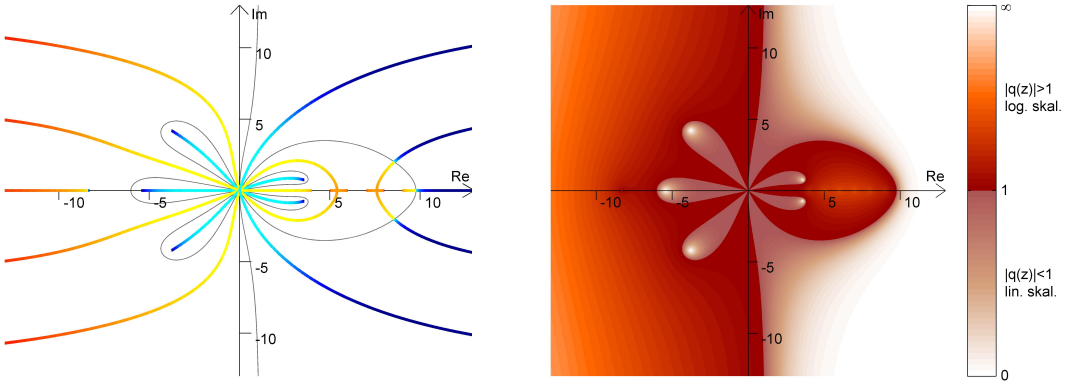

Abbildung 5.6: Ordnungspfeile und Ordnungsstern zu Gleichung (5.11)

Ordnungspfeile und Ordnungsstern der Funktion sind in Abbildung 5.6 dargestellt. Sie hat die Ordnung 6. Die Aufwärtspfeile um 0 haben die Winkel $\frac{2k\pi}{7}$ für $k = 0, 1, ...6$ inne. Der dritte Aufwärtspfeil verlässt 0 also mit einem Winkel von $\frac{4\pi}{7} \neq \frac{\pi}{2}$, er ist also nicht tangential zur imaginären Achse. Außerdem befinden sich alle Pole in der

rechten Halbebene. Aus dem Ordnungssternbild der Funktion kann jedoch deutlich abgelesen werden, dass sie nicht A-stabil ist. Dies ist ein großer Vorteil des Kriteriums zur A-Stabilität der Ordnungssterne gegenüber den Ordnungspfeilen.

5.2 Besondere Ordnungspfeillinien

Wie nach Definition 5.1 erwähnt, bestehen Ordnungsnetze nie nur aus dem Hauptordnungsnetz. Zusätzlich zu den in allen Ordnungspfeilbildern vorhandenen ungefähr zur x-Achse parallel verlaufenden Linien gibt es aber in manchen Ordnungspfeilbildern auch noch weitere Linien.

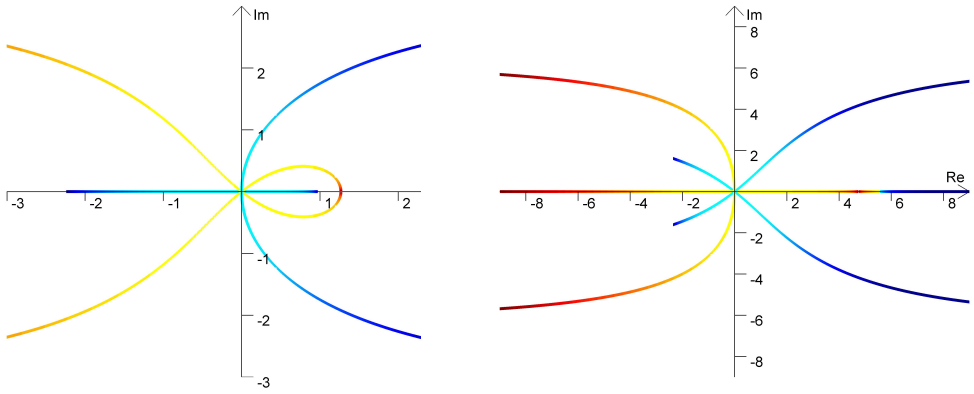

Abbildung 5.7: Ordnungspfeile der SDIRK3 Approximationen

Dies ist zum Beispiel häufig bei doppelten Polstellen der Fall. Als Beispiel werden wieder SDIRK3-Verfahren verwendet, hier in Abbildung 5.7. Für positives Vorzeichen in der Wurzel sieht man, dass zwei Aufwärtspfeile vom Ursprung zu dem zweifachen Pol führen. Für negatives Vorzeichen ergibt sich das rechte Bild, indem nur ein Pfeil zum Pol führt. Für den zweiten Pol muss es dann auch einen zweiten Pfeil geben, der von dem Pol wegführt. Da er nicht zum Ursprung führt, muss er als Abwärtspfeil im weiteren Sinne (er geht von ∞ über 1 weiter abwärts) zu $+\infty$ führen.

Für Funktionen mit mehreren reellen Polstellen kann als Beispiel die von Butcher vorgestellte Funktion in (5.11) betrachtet werden. Ihre Ordnungspfeile sind in Abbildung 5.6 links dargestellt. Die weiter vorne liegenden Polstellen sind hier zum Teil miteinander durch ab- und wieder aufwärts schwingende Pfeile verbunden und zu diesen Pfeilen führen Pfeile vom Ursprung, die zwischen den zwei Polen auf die x-Achse treffen.

Weiterhin gilt für die in Abschnitt 4.2 für Ordnungssterne vorgestellten Finger, die nicht mit dem Ursprung verbunden sind, dass auch Ordnungspfeile sie nicht mit dem Ursprung verbunden darstellen.
Dort wurde als Beispiel eine beschränkten Padé-Approximation von Nørsett verwendet. Ihr Ordnungspfeilbild ist hier in Abbildung 5.8 dargestellt. Es ist zu sehen, dass die

zwei Nullstellen durch zwei Ordnungspfeile verbunden sind, die jeweils vom gleichen Punkt auf der x-Achse aus abfallend sind.

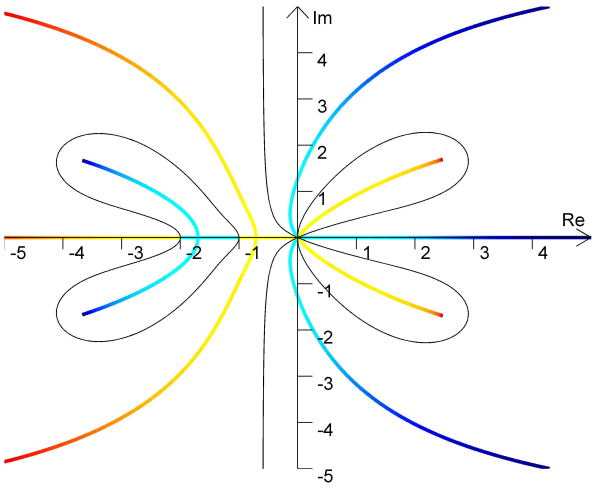

Abbildung 5.8: Ordnungspfeile der Funktion von Nørsett aus Gleichung (4.39)

Ordnungspfeile sind bei Butcher nicht immer sehr präzise definiert. Für Ordnungspfeile, die nicht direkt mit dem Ursprung verbunden sind, ist kein Anfangs- und kein Endpunkt festgelegt. Sie verbinden Polstellen mit $+\infty$ oder mit Nullstellen und Nullstellen mit $-\infty$ über einen Punkt, an dem der Betrag $|q(z)| = 1$ ist, hinweg.
Ob sie als Abwärts- oder Aufwärtspfeile klassifiziert werden können und von welchem Punkt sie ausgehen, ist dabei nicht klar. So können zum Beispiel die zwei die Nullstellen verbindenden Pfeile in Abbildung 5.8 als Aufwärtspfeile von den Nullstellen zur x-Achse oder als Abwärtspfeile von der x-Achse zu den Nullstellen interpretiert werden.

Für Abbildung 5.9 wurde eine Funktion $R(z)$ mit zwei komplexen doppelten Polstellen und einem reellen Pol mit der Formel aus Gleichung (4.49) erstellt. Sie zeigt als Beispiel einen von der doppelten Polstelle ausgehenden Abwärtspfeil zu einer Nullstelle.
Außerdem bringt dieses Beispiel noch einmal die im Beweis von Satz 5.3 dargestellte Problematik von Sattelpunkten hervor. Bei Butcher wird nicht eindeutig definiert, welcher weiterführende Pfeil an einem Sattelpunkt einen dort auftreffenden Pfeil fortsetzen soll. Im Beispiel trifft sogar ein Abwärtspfeil des Hauptordnungsnetzes auf den Sattelpunkt, von dem aus zwei weiter abwärts führende Linien zu $+\infty$ führen und eine auf der x-Achse aufwärts gehende Linie zu einer Polstelle führt.

Letztendlich ist jedoch durch die Symmetrie zur x-Achse keine eindeutige Definition notwendig.

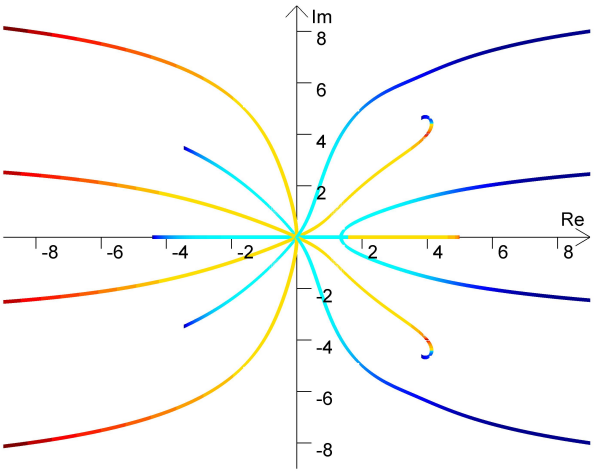

Abbildung 5.9: Ordnungspfeile Spezialfunktion - komplexe doppelte Polstelle

5.3 Beweis von Ehles Vermutung mit Hilfe von Ordnungspfeilen

Auch mit Ordnungspfeilen lässt sich die Vermutung von Ehle beweisen. Im Voraus wird nur ein neuer Satz und eine Folgerung daraus benötigt.
Der nächste Satz besagt, dass es mindestens so viele in Nullstellen und Polstellen endende Pfeile gibt, wie die Ordnung von $R(z)$ ist, bzw. dass die Ordnung maximal so groß ist, wie die Anzahl von Pfeilen die in Nullstellen und Polstellen enden.

Satz 5.5 *Sei* $R(z) = \frac{P(z)}{Q(z)}$ *eine Approximation an die Exponentialfunktion, für die* $\deg(P) = k$ *und* $\deg(Q) = j$. *Sei* k_1 *die Anzahl der Abwärtspfeile, die in Nullstellen enden und* j_1 *die Anzahl der Aufwärtspfeile, die in Polstellen enden. Dann gilt:*

$$p \leq k_1 + j_1 \tag{5.12}$$

Beweis: Es gibt nach Satz 5.3 $p + 1 - k_1$ Abwärtspfeile, die nicht in Nullstellen, sondern bei $+\infty$ enden und $p + 1 - j_1$ Aufwärtspfeile, die nicht in Polstellen sondern bei $-\infty$ enden.
Sei α der kleinste Winkel, der alle Abwärtspfeile oberhalb (und auf) der x-Achse, die bei $+\infty$ enden, einschließt und von der positiven x-Achse ausgehend gemessen wird. Wenn er an der x-Achse gespiegelt wird, sind somit innerhalb eines Winkels von 2α alle nicht zu Nullstellen führenden Abwärtspfeile eingeschlossen.
Analog sei der Winkel β der kleinste Winkel, der alle Aufwärtspfeile der oberen Halbebene, die bei $-\infty$ enden, einschließt, gemessen von der negativen x-Achse. Die Winkel werden in Bild 5.10 am Beispiel der Padé(2,3)-Approximation dargestellt.
Zwischen je zwei aufeinanderfolgenden Aufwärtspfeilen und zwischen je zwei aufeinanderfolgenden Abwärtspfeilen ist der Winkel $\frac{2\pi}{p+1}$ nach Satz 5.2. Der Zwischenraum zwischen $p + 1 - k_1$ bzw. $p + 1 - j_1$ Pfeilen, der jeweils durch 2α bzw. 2β ausgedrückt ist, umspannt also mindestens:

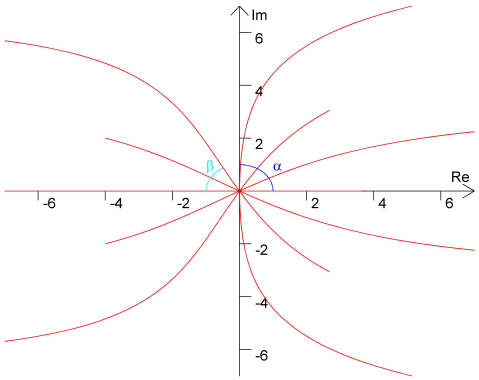

Abbildung 5.10: Ordnungspfeillinien mit Winkeln (Padé(2,3)-Approximation)

$$2\alpha \geq \frac{(p-k_1)2\pi}{p+1}, \quad \text{und} \quad 2\beta \geq \frac{(p-j_1)2\pi}{p+1} \tag{5.13}$$

Aufwärts- und Abwärtspfeile können sich nicht schneiden und es befindet sich mindestens ein Winkel $\frac{\pi}{p+1}$ zwischen dem letzten Abwärtspfeil der bei $+\infty$ endet und dem ersten Aufwärtspfeil, der bei $-\infty$ endet. Dieser Abstandswinkel wird mit γ bezeichnet. (Im Beispiel sind diese zwei Pfeile genau benachbart und der Abstand ist exakt der angegebene Bruch.) Addiert müssen diese drei Winkel genau einen Vollkreis ergeben und mit Einsetzen der Ungleichungen ergibt sich

$$2\pi = 2\alpha + 2\beta + 2\gamma \geq \frac{(p-k_1)2\pi + (p-j_1)2\pi + 2\pi}{p+1} \tag{5.14}$$

woraus durch Umformen folgt, dass $p \leq k_1 + j_1$. □

Bei den Padé-Approximationen gilt $p = k + j$ und mit der gleichen Umformung wie im Beweis von Satz 4.13 folgt, da $k_1 \leq k$ und $j_1 \leq j$, dass $k = k_1$ und $j = j_1$. Es ergibt sich also als

Folgerung 5.6 *Es enden bei Padé-Approximationen k Pfeile in Nullstellen und j Pfeile in Polstellen.*

Nun kann die Vermutung von Ehle bewiesen werden. Da sie nicht mehr nur eine Vermutung ist, wird sie bei Butcher auch als „Ehle-Grenze“ bezeichnet.

Satz 5.7 *Sei $R(z)$ die Padé(k,j)-Approximation an e^z. Dann ist das zugehörige Verfahren nur A-stabil, wenn $j \leq k + 2$.*

Beweis: Sei β gemessen von der positiven x-Achse der kleinste Winkel, sodass alle Aufwärtspfeile, die den Ursprung in der oberen Halbebene verlassen und in einem Pol enden, in β enthalten sind. Insgesamt enden mit der letzten Folgerung j Aufwärtspfeile in Polstellen. Der Abstand zwischen zwei aufeinanderfolgenden Aufwärtspfeilen ist $\frac{2\pi}{p+1}$ und somit unterscheiden sich zwei aufeinanderfolgende in Polstellen endende Pfeile

mindestens um diesen Abstand. Also gilt für den Winkel, den alle in Polstellen endende Aufwärtspfeile umspannen:

$$2\beta \geq \frac{2\pi(j-1)}{p+1} \tag{5.15}$$

Dieser Winkel muss kleiner sein als π, da sonst entweder Pole in der linken Halbebene liegen oder die Aufwärtspfeile über die imaginäre Achse hinüberreichen und damit R nicht A-stabil wäre. Also gilt mit $p = k + j$:

$$\frac{2\pi(j-1)}{k+j+1} < \pi \;\Leftrightarrow\; 2j-2 < k+j+1 \;\Leftrightarrow\; j < k+3 \;\Leftrightarrow\; j \leq k+2 \tag{5.16}$$

□

5.4 Ordnungspfeile bei Mehrschrittverfahren

Aus den in Abschnitt 4.6 beschriebenen einzelnen Blättern des Ordnungssterns können auch wieder Ordnungspfeile definiert werden, ähnlich wie bei den Einschrittverfahren als diejenigen Linien für die

$$q_j(z) = \frac{\zeta_j(z)}{e^z} \tag{5.17}$$

positiv und reell ist und das Argument 0 ist. Es sind in Abbildung 5.11 wieder die Blätter des BDF2-Verfahrens zu sehen.

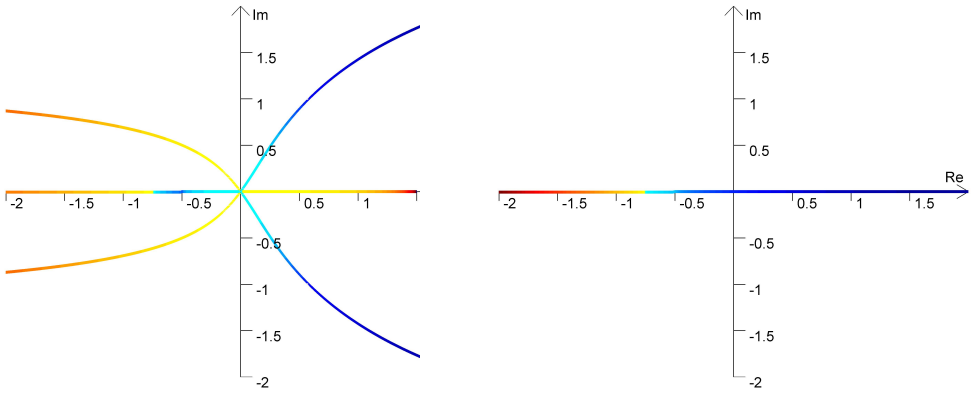

Abbildung 5.11: Ordnungspfeilblätter des BDF2-Verfahrens. Links ζ_1, rechts ζ_2

Mit den Ordnungspfeilen für Mehrschrittverfahren lässt sich auch die zweite Dahlquist-Schranke und die Daniel-Moore-Vermutung zeigen. Außerdem verwendet Butcher sie, um die von ihm und Chipman verfasste "Butcher-Chipman-Conjecture" zu beweisen, ein Beweis, der ohne Ordnungspfeile bis heute nicht gelungen ist. Es handelt sich um ein Analogon zur Ehle-Vermutung für verallgemeinerte Padé-Approximationen bei Mehrschrittverfahren. Erklärung und Beweis sind z.B. in "Order and Stability of Generalized Padé-Approximations" veröffentlicht 2008 von John Butcher in Science Direct zu finden.[5]

6 Kurzzusammenfassung

In dieser Arbeit wurden einige numerische Verfahren zum Lösen von gewöhnlichen Differentialgleichungen und die Bedeutung der Begriffe der steifen Differentialgleichung und der A-Stabilität beschrieben. Weiterhin wurden Padé-Approximationen vorgestellt und die Ordnung eines numerischen Verfahrens und einer Approximation an die Exponentialfunktion definiert.

Auf diesen Begriffen aufbauend wurden sowohl Ordnungssterne als auch Ordnungspfeile eingeführt und ihre Eigenschaften bezüglich der Ordnung und Stabilität beschrieben und ausführlich bewiesen. Dabei ging es hauptsächlich um Einschrittverfahren.

Es wurden Beweise von Sätzen, in denen es um die Ordnung und Stabilität von Einschrittverfahren geht, wie bei der Vermutung von Ehle, mit Hilfe von Ordnungssternen geführt. Außerdem wurde mit Ordnungssternen der Satz zum Vergleich von Stabilitätsbereichen und der Satz zur Beschränkung der Ordnung durch reelle Polstellen gezeigt. Der Beweis von Ehles Vermutung wurde auch mit Hilfe von Ordnungspfeilen ausgearbeitet.

Die Literatur von den Autoren Hairer, Wanner, Nørsett und Butcher, die der Arbeit als wichtige Quellen zugrunde lag, beschäftigt sich auch intensiv mit Mehrschrittverfahren. Hier wurde jeweils nur ein kurzer Ausblick auf Ordnungssterne bzw. Ordnungspfeile bei Mehrschrittverfahren mit einem Beispiel gegeben. Auch auf eine Behandlung der Ordnungssterne im Gebiet der Approximationstheorie wurde mit einem Beispiel kurz hingewiesen.

Literaturverzeichnis

[1] A. Abdi, J.C. Butcher: „Experiments with Order Arrows“, *SciCADE (International Conference on Scientific Computation and Differential Equations), Toronto, Juli 2011*

[2] M. Aigner, G.M. Ziegler: „Proofs from THE BOOK“, *Kapitel 10:* „Three applications of Euler's formula“, *Springer Verlag, 2008*

[3] L. Brugnano, F. Mazzia, D. Trigiante: „Fifty Years of Stiffness“, *in: „Recent Advances in Computational and Applied Mathematics“, Springer Verlag, 2011, S.1-21*

[4] J.C. Butcher: „Numerical Methods for Ordinary Differential Equations“, *John Wiley & Sons Ltd, 2008*

[5] J.C. Butcher: „Order and stability of generalized Padé approximations“, *Applied Numerical Mathematics 59, 2009, S. 558 - 567*

[6] J.C. Butcher: „Order Stars and Order Arrows“, *Konferenz anlässlich Ernst Hairers 60. Geburtstag, Genf, Juni 2009*

[7] R. V. Churchill: „Complex Variables and Applications“, *McGraw-Hill, 1974*

[8] H. Fischer, H. Kaul: „Mathematik für Physiker“ - *Kapitel 7*: „Einführung in die Funktionentheorie“, *Springer Verlag, 2011*

[9] E. Hairer, S.P. Nørsett, G. Wanner: „Order Stars and Stability Theorems“, *BIT 18, 1978, S. 475 - 489*

[10] E. Hairer, S.P. Nørsett, G. Wanner: „Solving Ordinary Differential Equations I, Nonstiff Problems“, *Springer-Verlag, 1986*

[11] E. Hairer, G. Wanner: „Solving Ordinary Differential Equations II, Stiff and Differential-Algebraic Problems“, *Springer-Verlag, 1996*

[12] P. Henrici: „Applied and Computational Analysis, Volume I, Power Series-Integration-Conformal Mapping-Location of Zeros“, *John Wiley & Sons Inc, 1974*

[13] A. Iserles, S. P. Nørsett: „Order Stars“, *Chapman & Hall, 1991*

[14] E. M. Stein, R. Shakarchi: „ Princeton Lectures in Analysis II, Complex Analysis“, *Princeton University Press, 2003*

Anhang

Anhangsverzeichnis

A Inhalte der CD

Der Diplomarbeit ist eine CD beigefügt. Im Dokumentenordner „Matlab“ befinden sich die erstellten Programme, sie sind ausgerichtet auf die Matlab Version R2010a. Für einige Berechnungen wurde auch die „Symbolic Math Toolbox“ verwendet. In Tabelle A.1 werden die Namen der Programme und eine Kurzbeschreibung aufgelistet.

Dateiname	Beschreibung
bspf.m	Funktion, die für die Beispieldgl. (3.1) die Ergebnisse mit exp. oder imp Euler-Verfahren berechnet
SteifeDglBsp.m	Skript, das zur gleichen Beispielfunktion Fehlertabellen und Graphen ausgibt
pade.m	Funktion, gibt bei Eingabe von z, k und j die Werte der Padé(k, j)-Approximation an die Exponentialfunktion an den Punkten z an
stabAdams.m	Skript, zeichnet Stabilitätsgebiet für Adams-Bashforth-4
stability.m	Skript, stellt verschiedene Stabilitätsgebiete dar
OSgui.m	GUI, Auswahl verschiedener Stabilitätsfunktionen und Ausgabe des Stabilitätsgebiets, des Ordnungssterns und der Ordnungspfeile ist möglich
orderStar.m	Funktion, stellt Ordnungssterne dar, auch für relative Ordnungssterne, allgemeine Ordnungssterne und Quiver-Bilder (Winkel von q als Pfeil auf ∂A) verwendet
orderArrow.m	Funktion, stellt Ordnungspfeile dar, auch für Winkelkonturbilder und Mischbilder verwendet
bspBerKoeff.m	Skript, zeigt als Beispiel auf, wie die selbsterstellten Funktionen berechnet wurden
koordinatensystem.m	Funktion, stellt Koordinatensystem der aktuellen Grafik dar

Tabelle A.1: Erstellte Matlab Programme

Alle Abbildungen die in dieser Arbeit verwendet werden, wurden mit den aufgelisteten Programmen selbst erstellt.
Im Dokumentenordner „Abbildungen“ sind alle in der Diplomarbeit verwendeten und einige weitere erstellte Abbildungen zu finden. Sie werden in Tabelle A.2 aufgelistet. Dabei wird unter Bemerkungen, falls nicht aus dem Namen erkenntlich oder in der Diplomarbeit angegeben, vermerkt, an welcher Stelle die Koeffizienten der abgebildeten Stabilitätsfunktion zu finden sind. Es wird auf die Matlab-Datei „OSGUI.m“ hingewiesen, in der unter „popupmenu1_Callback“ bei unterschiedlichen Fällen der „switch“-Anweisung die Koeffizienten dieser Funktionen aufgelistet sind.
Außerdem wird, falls vorhanden, eine Seitenzahl in der Diplomarbeit, auf der die Abbildung zu finden ist, angegeben.

Tabelle A.2: Erstellte Abbildungen

Ordnungssterne

Dateiname (.jpg)	S.	Bemerkungen
OSappAnSin	44	
OSBDF2zeta1	47	
OSBDF2zeta2	47	
OSbutcher6pole	54	Stabilitätsfunktion in Glg. (5.11) auf S. 54
OsexpEuler		
OSexpEulerMitAsymp		
OSexpEulerMitAsympKl		
OSfunktionJ5K2O5	46	Koeff. in „OSgui.m“, „case 25“, $q = qq2$
OSfunktionJ5K3O5bsp1		Koeff. in „OSgui.m“, „case 25“
OSfunktionJ5K3O5bsp2		Koeff. in „OSgui.m“, „case 25“
OSfunktionJ5K3O6		Koeff. in „OSgui.m“, „case 25“
OSfunktionJ6K3O6		Koeff. in „OSgui.m“, „case 26“
OSfunktionJ6K4O7		Koeff. in „OSgui.m“, „case 26“
OSimpEuler		
OSimpEulerMitAsymp		
OSimpEulerMitAsympKl		
OSkomplexeDoppeltePoleI		Koeff. in „OSgui.m“, „case 17“
OSkomplexeDoppeltePoleII		Koeff. in „OSgui.m“, „case 18“
OSkomplexeDoppeltePoleIII		Koeff. in „OSgui.m“, „case 19“
OSkomplexeDoppeltePoleIV		Koeff. in „OSgui.m“, „case 20“
OSkomplexeDoppeltePoleV		Koeff. in „OSgui.m“, „case 21“
OSkomplexeDoppeltePoleVI		Koeff. in „OSgui.m“, „case 22“
OSkomplexeDoppeltePoleVII		Koeff. in „OSgui.m“, „case 23“
OSkomplexeDoppeltePoleVIII		Koeff. in „OSgui.m“, „case 24“
OSpade03	26	
OSpade12	26	
OSpade21	26	
OSpade30	26	

Tabelle A.2: Fortsetzung

Dateiname (.jpg)	S.	Bemerkungen
OSpade08		
OSpade17	35	
OSpade26	35	
OSpade35	35	
OSpade44	35	
OSpade53	23	
OSpade62	35	
OSpade71		
OSpade80		
OSpade23mitAsymp	30	
OSRK4		
OSRK4MitAsymp		
OSRK4MitAsympkl		
OSs22Norsett	34	
OSSdirk3gneg		
OSSdirk3gpos		

Ordnungssterne Winkel auf ∂A

Dateiname (.jpg)	S.	Bemerkungen
OSQuiverPade12		
OSQuivers22Norsett		
OSQuiverSDIRKgnegGr	32	
OSQuiverSDIRKgposGr	31	

Relative Ordnungssterne

Dateiname (.jpg)	S.	Bemerkungen
OSrelExpdurchImp	41	
OSrelImpdurchExp	41	
OSrelRK4durchRK3	41	
OSrelSDIRKnegdurchpos	41	

Ordnungspfeile

Dateiname (.jpg)	S.	Bemerkungen
OABDF2zeta1	59	
OABDF2zeta2	59	
OAExpEulerGr	50	
OAimpEuler		
OAkomplexeDoppeltePoleI		
OAkomplexeDoppeltePoleII		
OAkomplexeDoppeltePoleIII	57	
OAkomplexeDoppeltePoleIV		

Tabelle A.2: Fortsetzung

Dateiname (.jpg)	S.	Bemerkungen
OAkomplexeDoppeltePoleV		
OAPade23		
OApade23mitWinkeln	58	
OApade32	53	
OApade42	53	
OApade35	48	
OARK3		
OAsdirk3gneg	55	
OAsdirk3gpos	55	

Ordnungssterne und Ordnungspfeile Mischbilder

Dateiname (.jpg)	S.	Bemerkungen
OAOSbutcher6pole	54	
OAOSExpEulerGr		
OAOSpade23	49	
OAOSs22norsett	56	
OAOSsdirk3gneg		
OAOSsdirk3gpos		
OSOApade23		

Stabilitätsgebiete

Dateiname (.jpg)	S.	Bemerkungen
ExpEulerStability	16	
ImpEuStability	16	
RKStabilityExpl	16	
RKStabilityExplScaled2	43	
scaledStabRK1		
scaledStabRK2		
scaledStabRK3		
scaledStabRK4		
SDirknegStability	17	
SDirkposStability	17	
StabAdams	21	
StabPade04exaAlpha	20	

Winkelbilder

Dateiname (.jpg)	S.	Bemerkungen
WinkelKonturPade12	49	
WinkelPade12	49	

Tabelle A.2: Fortsetzung

Dateiname (.jpg)	S.	Bemerkungen
WKbutcher6pole		
WKRK3		

Sonstige Bilder

Dateiname (.jpg)	S.	Bemerkungen
bspffunktExpVsImpEuh1d49	13	
bspffunktExpVsImpEuh1d50	13	
graphOSpade25	38	
graphOSS22	38	

B Originalzitat aus [6]

Auf der Konferenz anlässlich Ernst Hairers 60. Geburtstag stellte Butcher in dem Vortrag „Order Stars and Order Arrows“ eben diese vor. Dabei benutzt er in den Folien folgende Worte um einige hier nicht beigefügte Bilder zu beschreiben:
„[...] Part one: The beautiful theory of A-stability, following its first announcement by Dahlquist, is like a beautiful garden in a Swedish summer, with the midnight sun clearly visible.
Part two: The garden is becoming popular as more mathematicians visit it to make their own contributions to A-stability.
Part three: Now the garden is so crowded and polluted with theorems, theses and more and more definitions, that it is no longer beautiful and even the midnight sun is hidden.
Part four: Order stars have been introduced and the garden is beautiful again. A new species of flower now replaces some of the weeds and the midnight sun is shining again.
[...]
During the long Swedish winter, it is too cold to go into the Gården. It might be better to stay indoors and play a board game such as Snakes and Ladders.
The order arrows “Snakes and Ladders” board game [...]“

Erklärung

Ich versichere, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe, insbesondere sind wörtliche oder sinngemäße Zitate als solche gekennzeichnet. Mir ist bekannt, dass Zuwiderhandlung auch nachträglich zur Aberkennung des Abschlusses führen kann.

Ort, Datum Unterschrift